Dissecting human cortical development, evolution and malformation using organoids and single-cell transcriptomics

Von der Fakultät für Lebenswissenschaften
der Universität Leipzig
genehmigte

DISSERTATION

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM

Dr. rer. nat.

vorgelegt
von Master of Science Sabina Kanton
geboren am 27.01.1990 in Zeitz

Dekan: Prof. Dr. Marc Schönwiesner

Gutachter: Prof. Dr. Svante Pääbo
Prof. Dr. Wolfgang Enard

Tag der Verteidigung: 05.06.2020

THIS THESIS IS BASED ON THE FOLLOWING PUBLICATIONS:

1. **Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development**

   Felipe Mora-Bermúdez*, Farhath Badsha*, Sabina Kanton*, J. Gray Camp*, Benjamin Vernot, Kathrin Köhler, Birger Voigt, Keisuke Okita, Tomislav Maricic, Zhisong He, Robert Lachmann, Svante Pääbo, Barbara Treutlein, Wieland B. Huttner

   eLife 26(5) (2016)

2. **Organoid single-cell genomic atlas uncovers human-specific features of brain development**

   Sabina Kanton*, Michael James Boyle*, Zhisong He*, Malgorzata Santel, Anne Weigert, Fatima Sanchis Calleja, Patricia Guijarro, Leila Sidow, Jonas S. Fleck, Dingding Han, Zhengzong Qian, Michael Heide, Wieland B. Huttner, Philipp Khaitovich, Svante Pääbo, Barbara Treutlein and J. Gray Camp

   Nature 574(7778) (2019)

3. **Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia**

   Johannes Klaus*, Sabina Kanton*, Christina Kyrousi*, Ane Cristina Ayo-Martin, Rossella Di Giaimo, Stephan Riesenberg, Adam C. O'Neill, J.Gray Camp, Chiara Tocco, Malgorzata Santel, Ejona Rusha, Micha Drukker, Mariana Schroeder, Magdalena Götz, Stephen P. Robertson, Barbara Treutlein and Silvia Cappello

   Nature Medicine 25(4) (2019)

* shared first authors

BIBLIOGRAPHISCHE DARSTELLUNG

Sabina Kanton
Dissecting human cortical development, evolution and malformation using organoids and single-cell transcriptomics
Fakultät für Lebenswissenschaften
Universität Leipzig
Kumulative Dissertation
196 Seiten, 331 Literaturangaben, 43 Abbildungen, 31 Tabellen

---

During the last years, important progress has been made in modeling early brain development using 3-dimensional in vitro systems, so-called cerebral organoids. These can be grown from pluripotent stem cells of different species such as our closest living relatives, the chimpanzees and from patients carrying disease mutations that affect brain development. This offers the possibility to study uniquely human features of brain development as well as to identify gene networks altered in neurological diseases. Profiling the transcriptional landscape of cells provides insights into how gene expression programs have changed during evolution and are affected by disease. Previously, studies of this kind were realized using bulk RNA-sequencing, essentially measuring ensemble signals of genes across potentially heterogeneous populations and thus obscured subtle changes with respect to transient cell states or cellular subtypes. However, remarkable advances during the last years have enabled researchers to profile the transcriptomes of single cells in high throughput.

This thesis demonstrates how single-cell transcriptomics can be used to dissect human-specific features of the developing and adult brain as well as cellular subpopulations dysregulated in a malformation of the cortex.

In chapter 1, cerebral organoids generated from human and chimpanzee induced pluripotent stem cells are used to identify similarities and differences in neural progenitor cells and neurons of the developing cortex using single-cell RNA-sequencing. Many features of cortex development are remarkably similar between human and chimpanzee, however, cell type-specific gene expression differences at the level of progenitor cells and neurons are detected, thus showing the feasibility of applying organoids for modeling evolutionary differences in controlled culture environments.

Chapter 2 presents a temporally resolved atlas of organoid development using an extended set of individuals, species, time points and cells compared to the data set presented in chapter 1. Moreover, gene expression changes are integrated with regulatory information from single-cell chromatin accessibility measurements to dissect potential regulatory mechanisms underlying gene expression differences between species. In addition, single-nucleus transcriptomes from human, chimpanzee/bonobo and macaque adult prefrontal cortex tissue are analyzed to identify species-specific expression differences that can be detected only in adult stages or such that are commnon to both the adult and the developing cortex.

Chapter 3 presents a comparative analysis of cerebral organoids derived from patients with cortical malformations (e.g. periventricular heterotopia, PH) and control organoids, showing that PH-related malformations can be modeled in the cerebral organoid system. Single-cell RNA-seq analyses on cerebral organoids and 2-dimensional neuronal cultures reveal subpopulations of cells displaying dysregulated gene expression patterns and migratory behavior, suggesting that an altered navigation system in a subset of neurons underlie this form of PH.

Table of contents

Summary

Introduction

Humans live in intricately organized societies and have developed technologies that allow them to cope with diverse habitats and living conditions. This can only be achieved through a high level of cooperation, communication via language and complex planning, that allow for adaptation and knowledge transmission across multiple generations [1]. Although our closest living relatives, the great apes, and more specifically chimpanzees, show higher order cognitive abilities such as tool use and even behaviors that have been associated with tradition and culture [2, 3], humans are unique in the extent of their technological advancements and the way they have shaped the environment. These cognitive differences between humans and chimpanzees have a basis in the form, function and cellular composition of the brain, and these differences likely have a root during early developmental phases. However, to date, little is known about how brain development differs between humans and great apes and it has been difficult to address mechanisms underlying potential differences due to the lack of a tractable system that effectively models great ape brain development.

Brain development and human-specific brain characteristics

Early phases of brain development

The brain develops from the neuroectoderm which forms the neural plate and folds into the neural tube after closure. The neural tube then divides into three vesicles, the forebrain (prosencephalon), midbrain (mesencephalon) and hindbrain (rhombencephalon) with the cerebral cortex arising from the prosencephalon which is subsequently termed the telencephalon. The neural tube consists of a monolayer of cells known as the neuroepithelium which harbors stem cells that are oriented along an apico basal axis from the lumen (apical side) of the tube to the outer side (basal side). Neuroepithelial cells are the stem cells of the developing brain and divide and give rise to radial glial cells with restricted fate potential that in turn differentiate into different progenitor cell types [4].

The various cortical progenitor cells differ in their proliferative and differentiation potential, location of mitosis and contact to the apical and basal lamina [5] and are located in different germinal zones of the developing cortex. The most apical zone is called the ventricular zone (VZ), followed by the subventricular zone (SVZ), intermediate zone, subplate, cortical plate and the so-called marginal zone. During neurogenesis in the cortex, excitatory glutamatergic neurons arise from the progenitor cells and migrate along the radial glia fibers in an inside-out fashion to their location at the basal side of the cortex where they eventually constitute six cortical layers characterized by distinct cell composition [6–9]. Other cell types, such as the majority of GABAergic inhibitory neurons (interneurons) in the cortex, develop from distinct germinal zones, called the ganglionic eminences and migrate into the neocortex to integrate into neuronal circuits [10–12].

Later during development, glia cells such as astrocytes and myelinating oligodendrocytes develop from the radial glia cells. Moreover, neuronal circuits are established and refined by synapse formation. Brain development is not completed upon birth, many processes such as brain growth, synaptogenesis, gliogenesis and myelination extend beyond fetal development into postnatal stages and for example myelination even still undergoes changes in the adult brain [4, 13]. The adult brain eventually harbors around 86 billion neurons [14] with 10-20 billion neurons and 15-30 billion glial cells located in the cerebral cortex [15] that accounts for more than 80% of the total brain mass [16]

Aberrations during cortical development lead to variable malformations which can be caused by e.g. impairments in progenitor behavior, neuronal migration and circuit integration. As a result, alterations in brain size (microcephaly, macrocephaly), folding of the cortex (e.g. lissencephaly) and mislocalization of neurons (heterotopia) [17] can be observed. Many of these conditions lead to epileptic seizures and cognitive disability. Moreover, it has been proposed that even neuropsychiatric diseases such as autism [18–21] and schizophrenia [19] have a developmental component and might be rooted in human-specific features of brain development.

Human-specific features of the developing and adult brain

One of the major characteristics distinguishing the human brain from that of our closest living relatives, the great apes, is the enlargement of the neocortex. Consequently, humans display an approximately three fold increase in brain mass compared to chimpanzees [22]. This enlargement could be due to cell cycle or developmental timing differences, or changes in the characteristics of neural stem and progenitor cells. For example, the neurogenic period in humans is longer than in macaques with 100 days compared to 60 days [23]. Moreover, this enlargement has in part been attributed to the SVZ which is thickened in primates compared to mice and further subdivided in an inner (ISVZ) and outer portion (OSVZ) which is delimited by an inner fiber layer [24]. The OSVZ harbors a special cell type called basal radial glia (also termed outer radial glia) with increased proliferative capacity and therefore neuronal output thus potentially contributing to increased brain size [25, 26]. Basal radial glia are nearly absent in mice [27], less prominent in ferrets [26] and macaque [28] and are marked by specific transcriptional profiles [29], morphological variability [30], complex lineage trajectories [28] and the retainment of stem-like and proliferative features by generating their own proliferative niche [29, 31]. In addition to the overall increase in size of the human brain, it is still a matter of debate whether some brain regions, such as the prefrontal cortex, responsible for higher order processing, planning and social abilities [32], underwent a disproportionate enlargement in humans [33–38].

Aside from the above mentioned brain size difference, other morphological, cellular, physiological and molecular distinctions of the human brain have been described. For example, magnetic resonance imaging of human and chimpanzee brains revealed increased plasticity in cortical organization and morphology in humans and hypothesized that this could contribute to human-specific learning capabilities [39]. In the frontal pole, pyramidal neurons in layer III of the cortex are more widely spaced in humans than in apes [40] and generally show more complex morphology [41]. Moreover, humans are characterized by an enlargement of the superficial cortical layers (II+III) which could lead to changes in signaling between cortical regions [42].

Some human-specific brain differences might be caused by cell type innovations or enrichment of cell types. Van Economo neurons, a pyramidal cell type that can be found in many mammals with large brains, for example occurs more frequently in humans than in other great apes and is characterized by a bigger cell body in humans [43–45]. Profiling gene expression in different regions of the human, chimpanzee and macaque brain identified an interneuron subtype which also expressed dopamine synthesis markers [46]. This cell type was more prominent in the human striatum and only present in the neocortex of humans compared to great apes [46]. Another recently identified human interneuron cell type, designated 'rosehip cells', located in layer I of the cortex, was discovered based on morphological and gene expression analysis. However, it is not yet clear if this cell type is unique to humans since the study compared only humans and mice [47]. Astrocyte morphology and function also have diverged, for example so-called varicose projection astrocytes in layer V/VI which are absent in mice and lower primates but can be found in human and chimpanzee, although in the chimpanzee with decreased size and complexity [48]. Furthermore, humans have developed an increased glia/neuron ratio across different areas of the frontal cortex [49].

Comparative investigations into the composition of metabolites [50, 51] and lipids [52, 53] have shed light on dynamic and species-specific differences during brain development. For example, changes in glutamate metabolism have been detected in human prefrontal cortex but not cerebellum when comparing human, chimpanzee and macaque tissues [50]. In addition, there are generally more metabolic changes in the prefrontal cortex in humans compared to chimpanzee, macaque and mice [51]. Similar findings have been made with respect to lipids where more changes were detected in the neocortex compared to the cerebellum [52].

Human brain development appears to be slowed down, compared to that of chimpanzees and macaques, a phenomenon known as neoteny. It has been hypothesized that the elongated phases of development might allow for more flexibility and extended phases of learning needed to navigate in complex social constructs [54]. Gene expression studies have revealed subsets of genes showing similar timing shifts reminiscent of neoteny highlighting genes involved in synapse development [55–57]. Moreover, in humans a process called synaptic pruning is prolonged and also myelination takes longer [13, 58]. Similar timing differences might also play a role during prenatal development as modeled in in vitro systems [59] but these differences remain poorly understood.

Potential genetic underpinnings of human-specific brain features

The genetic underpinnings of human-specific features of the brain are not well understood yet and only some candidates have been functionally characterized in more detail. For example, human-specific deletions [60] and gene duplications [61] have been identified to be associated with evolutionary brain differences. The human-specific segmental duplication *ARHGAP11B* [62] has been shown to increase basal progenitor amplification and causes folding of the cortex in mice which points to a potential role in the evolution of brain size in humans. One more prominent example is *NOTCH2NL* [63, 64] which is expressed in radial glia cells in the fetal human brain and is involved in prolonged radial glia proliferation and could thus lead to brain size increase. *SRGAP2C*, another segmental duplication, induces a delay in dendritic spine maturation and leads to higher spine density when tested in the mouse neocortex [65, 66]. A candidate gene present in humans and great apes, *GLUD2* [67], could play a role in lipid biosynthesis during postnatal brain development.

Conserved sequences being deleted in humans (hCONDELs) might also contribute to evolutionary changes. One of those human-specific conserved deletions abrogates a putative enhancer of the gene *GADD45G* and could be responsible for increased brain size due to its role in tissue proliferation. Generally, human conserved deletions seem to be overrepresented near genes involved in neurogenesis of the cortex and often link to genes likely involved in increased proliferation [60].

The most distinguished example at the level of protein coding sequence differences to date is *FOXP2*, which is characterized by two amino acid changes specifically on the human lineage and has been attributed to differences in vocalization and learning when humanized in mice [68, 69]. Studies in mouse embryos electroporated with the human *FOXP2* version suggested that it might also be involved in early development by changing the relative proportion of radial glia to more differentiated neural progenitors [70].

GPR56, a gene involved in progenitor proliferation and cortical folding, exhibits increased number of promoters compared to mice and the deletion of one of these regulatory elements led to regionally restricted folding differences in the human brain, suggesting specification of cortex folding patterns by regulatory diversification [71].

Human accelerated regions (HARs) are genomic segments that exhibit an excess of sequence changes on the human lineage but are conserved in other species [72]. Generally, many of these accelerated regions are located in non-coding genomic regions and have been linked to enhancers active in brain development [73]. The gene which has the most HARs in its proximity is *NPAS3* and it was shown in mice that the human version of one of these HARs, 2xHAR142, leads to a broader activity pattern in humans that also includes the forebrain compared to the mouse and chimpanzee version [74]. Another interesting example represents human accelerated region 5 (*HARE5*) which controls gene expression of *FZD8*. When tested in mice, the human version leads to a faster cell cycle in neural progenitors and an increase in brain size compared to the chimpanzee sequence [75]. Human accelerated region 1 is one more prominent example, which accumulated 18 substitutions on the human lineage and is located within a regulatory RNA gene, a long noncoding RNA (lncRNA), *HAR1F*. The regulatory RNA is expressed in the cortex during early development in so called Cajal-Retzius cells which play an important role in establishing the laminar organization of the cortex and it is a rare example of a regulatory RNA that could play a role in the evolution of human brain development [72].

The above described studies have aided in linking human-specific genetic changes to evolutionary differences in brain development. However, most of them were conducted in mice and not in the endogenous human genetic context. Thus it would be desirable to use model systems that more faithfully recapitulate human biology.

Advanced in vitro model systems to study development

Previous models and the advent of induced stem cells

It has been hypothesized that many of the changes that can be observed in the human brain are rooted in early development, thus highlighting the significance of comparative developmental investigations. Studying human brain development, its malformations and evolution has mainly relied on using either fetal tissue [76] or model organisms such as mice [8], rats [6], ferrets [77] and macaques [57]. However, the use of fetal tissue presents ethical issues and technical limitations and it is impossible to obtain fetal tissues from great apes. Moreover, primary tissue cannot be easily manipulated via genetic tools to study the effect of selected genes. Transgenic model organisms such as mice on the other hand have often only partially been able to mirror the phenotypes seen in humans [78, 79] most likely due to their phylogenetic distance. Thus, new model systems are required that recapitulate human development in a more faithful manner to study evolution and disease.

A major breakthrough towards this goal was the discovery that somatic cells can be reprogrammed into pluripotent cells by administering certain factors [80] on comparatively easily accessible cell sources such as fibroblasts [80] or blood cells [81]. This has made it possible to generate stem cells from different primate species [82–84] and patient-specific stem cells to study diseases [85]. Due to their differentiation capacity, these cells can be coaxed into virtually any cell type and offer the potential to study different tissues from multiple individuals and species across multiple stages of development to decipher temporally resolved developmental programs. To investigate human-specific differences of brain development, a set of differentiation protocols in 2D have been developed [86–88]. Although these systems provide only a limited amount of complexity compared to the native tissue context, they for instance have proven useful in revealing differences in developmental timing across primate species [59, 89].

Organoids as state-of-the-art 3-dimensional in vitro models

During the last years, so-called organoid technologies have been introduced that allow modelling early stages of development using adult, embryonic or induced pluripotent stem (iPS) cells and thereby revolutionized the field of stem cell engineering. Organoids are 3-dimensional (3D) models that self organize into organ-like structures with an increased degree of complexity and recapitulate some functionality of the tissue including the appearance of tissue-specific cell types [90]. Organoids can for instance be generated from induced pluripotent stem cells or embryonic stem cells from humans. A plethora of different tissues has been modeled to date such as intestine [91], brain [92–94], liver [95], kidney [96], lung [97], pancreas [98], optic cup [99] and inner ear [100]. One of the key advantages of using stem-cell derived organoids over primary tissue is their accessibility to genetic manipulation by using e.g. CRISPR/Cas9 [101, 102] to study the effect of specific genes and variants as well as the 3D culture environment that more faithfully mimics developmental processes.

Types of brain organoids

Brain organoids generated from human embryonic stem cells have been introduced in 2008 by the Sasai group [103] with later modifications by Lancaster *et al.* [92] and Kadoshima *et al.* [93]. The system introduced by Lancaster *et al.* produces different regions of the brain in one organoid, which also led to the term 'cerebral organoids' for this system. Alternative methods are optimized towards coaxing cells into specific regional identities such as forebrain [93, 104], midbrain [104, 105] and hypothalamus [104], choroid plexus and hippocampus [106], hindbrain [107] and pituitary [108]. Numerous modifications and further enhancements of these protocols have been introduced including miniaturized organoids that are grown in 3D-printed bioreactors and so called assembloids to study the interaction between different brain regions [109–112]. Investigations of histological [92, 104, 113], transcriptomic [20, 114–

117], epigenomic [115, 116] and electrical network activity features [118] have corroborated the feasibility of organoids to study developmental processes reminiscent of early to mid-fetal development in humans.

Applications and prospects

Brain organoids were utilized to study neurodevelopmental disorders, such as microcephaly [92], cortical folding differences [102, 119] or neuronal migration defects [109, 120], Down Syndrome [121] as well as neuropsychiatric diseases with a developmental component such as autism [122]. For example, organoids were applied to model Miller-Dieker Syndrome, a form of lissencephaly leading to absence of folding of the cortex and reduced brain size. The authors detected differences in outer radial glia cell division and increased cell death in neural stem cells [119]. Moreover, organoids were utilized to elucidate mechanisms of Zika virus (ZIKV) infection [104, 123, 124], to test ZIKV infection inhibitors [125], to study the effect of bisphenol A on neurodevelopment [104] and even to model neurodegenerative diseases such as Alzheimer's disease [126, 127]. Hence, brain organoids have already proven valuable in tackling diverse questions related to brain development and disease.

Current drawbacks of organoids, however, encompass a limited maturation capacity with premature layering of cortical regions [104] as well as limitations in size due to the absence of external signaling cues and nutrient supply. Furthermore, the majority of protocols is marked by variability in composition, in particular, regarding the abundance of cells with specific brain regional identities, as well as from batch to batch [113, 128]. Thus, recent efforts focusing on improving the organoid system are geared towards increasing reproducibility [117, 129], controlling regionalization [104, 117, 129], introducing currently missing cell lineages such as oligodendrocytes [130, 131] and microglia [132] and improving maturation by extended cultivation times [133], supplementing additional growth factors [104, 128] or introducing vasculature [134] to expand the applicability of organoids even further.

Studying early brain development in humans and great apes has been a challenge due to ethical and technical limitations. Cerebral organoids that can be generated from pluripotent stem cells offer the possibility to model early developmental differences between those species in controlled cell culture environments. Previous studies [20, 59] have already shown the potential to study differences that distinguish early human brain development from that of great apes, such as a higher activation of the mTOR pathway in human radial glia compared to chimpanzee [20].

High throughput sequencing studies of brain development

Gene expression studies

Humans and chimpanzees share 98.8% of their DNA sequence, however phenotypic distinctions are apparent despite this high similarity [135]. It has been proposed that many differences that set humans apart from chimpanzees are due to noncoding regulatory sequence differences and changes in gene expression rather than sequence differences in protein-coding genes [136, 137]. Thus, profiling genome-wide patterns of gene expression and regulatory signatures are promising approaches to dissect global changes during brain development and in the adult brain that set humans apart from great apes.

Transcriptome sequencing efforts in the human brain have revealed changes during prenatal [76, 138] and postnatal brain development [138] and described dynamic gene expression changes with an excess of differences detected in the prenatal period [138]. In studies investigating postnatal stages of brain development, human developmental patterns seemed to be more diverged especially in the prefrontal cortex [139] and synapse-related gene expression in humans was characterized by a shift towards later time points in development compared to chimpanzee and macaque [56]. Similarly, investigating the prenatal and postnatal macaque brain in comparison to humans [57] revealed a longer maturation in the human brain linked to genes playing a role in myelination and synaptogenesis [140].

By dissecting layer-resolved gene expression patterns in prefrontal cortex of human, chimpanzee and macaque, differences indicating reorganization of layer expression patterns were detected and were observed more frequently in humans than chimpanzee suggesting cortical reorganization on the human lineage [141]. Moreover, the construction of gene coexpression networks offers an integrated view on gene relationships [142–145] and has revealed transcription factors such as Krüppel-associated box-zinc finger proteins (KRAB ZNFs) to be overrepresented among transcription factors with gene expression changes between human and chimpanzee in prefrontal cortex [143] and identified human-specific gene expression modules in frontal pole [144].

Nevertheless, in a broader context, it is not yet fully resolved if the human brain underwent accelerated evolution indicating an excess of changes on the human lineage or if the changes accumulated linearly over evolutionary divergence time [146–150]. Moreover, it is not well resolved yet if particular cell types underwent accelerated evolution.

Gene regulatory differences

A number of studies investigated evolutionary changes in transcriptional regulation such as DNA methylation, histone modifications or expression of regulatory RNAs such as micro RNAs and lncRNAs. For example, a comparison of histone modifications between human, macaque and mouse in prenatal developing brain found that regions of activity gains specific to humans were among others connected to cell proliferation and extracellular matrix [151]. A similar study conducted in adult brain of human, chimpanzee and macaque reported a small set of human-specific active regions with 139 enhancers that gained activity and 45 that lost activity in humans [152]. Comparable results were obtained when active transcription start sites in neurons of the prefrontal cortex were examined revealing 410 loci that gained and 61 loci that lost activity specifically in humans [153]. A recent study showcased the use of iPSC-derived neural crest cells to study epigenomic differences between human and chimpanzee that can explain craniofacial divergence, highlighting the versatility of stem cell models to study evolutionary differences [154]. Regulatory RNAs such as microRNAs [46, 139, 155] or lncRNAs [156] have only started to be explored in the context of human-specific brain differences. DNA methylation might also play a role in gene expression differences as was indicated by comparing prefrontal cortex of human and chimpanzee with higher methylation levels detected in chimpanzee [157].

Although these studies present valuable resources to dissect human-specific gene expression differences and their potentially underlying regulatory logic, the majority relies on bulk measurements. Thus, some of the changes observed might be due to compositional differences in the tissue and more subtle cell type-specific changes are obscured by ensemble averaging.

The landscape of single-cell technologies

Technological advances in the last years have made it possible to study the gene expression profiles of single cells [158] and single-cell analysis has become a widely used method due to its power to resolve complex tissue compositions, detect rare cell types and compare orthologous cell types across different species [159] at unprecedented resolution.

The first single-cell transcriptome methods were based on manual cell isolation and low throughput library generation in tubes [158]. An increase in throughput was achieved by valve-based microfluidic devices geared towards separating and isolating cells (Fluidigm C_1 system) [160–162], cell sorting and automatic handling in low volumes (MARS-seq, massively parallel RNA single-cell sequencing) [163] or using micropatterned plates with nanowells for cell isolation [164]. The introduction of droplet-based microfluidic techniques that rely on separating cells into water-in-oil droplets as distinct reaction containers has however substantially advanced the throughput of single-cell studies, making it amenable to profile thousands of cells in one experiment [165, 166]. Additional efforts, so called split and pool approaches, have enlarged the amount of cells that can be profiled in parallel even further allowing for the detection of rare cell types [167].

These single-cell transcriptomic techniques already had a substantial impact on the study of differentiation dynamics [161, 168], cell state transitions [169] cellular composition of tissues [170–173], lineages [174, 175] and effects of perturbations [176, 177]. Recent collaborative efforts have used single-cell transcriptomics to profile the cellular composition of all major organs of the mouse [178] as well as of humans [179] in order to create cell atlases in normal [163, 169, 180–184] and disease conditions [184–188]. Further, cellular reference maps were also generated for commonly [189–192] and less commonly used model organisms [168, 193]. Single-cell analysis has expanded from studying transcriptomes [194, 195] to genomes [196], epigenomes [197, 198], proteomes [199], chromosome conformations [200, 201] and various combinations thereof [202–207] leading to an integrated view on single cell states.

Studying the brain at single cell resolution

In the neuroscience field, single-cell transcriptomics has evolved into an invaluable tool to dissect brain cell type diversity in the mouse [208–211]. Approaches have been extended to combine single-cell sequencing with additional modalities such as cell morphology and electrophysiological properties [212, 213] as well as to interrogate gene expression patterns from frozen and archived tissue [214–216]. Single-cell transcriptome analysis has been applied to shed light on mechanisms of the developing [114, 217–220] and adult [215, 218, 221] human brain including the characterization of 3D cell culture models both from human [20, 114, 117, 128] as well as non-human primates [20, 222]. Finally, a comparison between mouse and human identified patterns of conservation and species-specific innovations [159].

Aim of the thesis and outline

The availability of methods to recapitulate human brain development in controlled culture environments, together with technologies to dissect tissues with single-cell transcriptomics, enable exciting opportunities to understand the mechanisms underlying human brain evolution and disease. In this thesis, stem cell-derived cerebral organoids and single-cell transcriptomics methods are used to explore human-specific features of neocortical development and adult prefrontal cortex cell diversity, as well as to investigate aberrant cell populations in a cortical malformation.

The first two chapters focus on using cerebral organoids to understand human-specific brain development. In chapter 1, cerebral organoids generated from iPS cells of human and chimpanzee are used to model early developmental changes on the human lineage using single-cell transcriptomics to decipher cell type-specific differences that likely contribute to human-specific cell type signatures and behavior.

Chapter 2 expands on the project described in chapter 1 by analysing single-cell transcriptomes from an increased number of cells, individuals, species and time points to study dynamic differences occurring during early brain development as recapitulated in cerebral organoids. The observed gene expression changes are further integrated with differences in chromatin accessibility between the species as well as with differential gene expression analysis performed on single-nucleus transcriptomic data of adult brain tissue from humans, chimpanzee/bonobos and macaques.

In chapter 3, cerebral organoids are used to understand cell states that emerge during a form of cortical malformation. Cerebral organoids generated from patient-derived iPS cells are applied to model periventricular neuronal heterotopia (PH), which is caused by mutations in the cadherin receptor-ligand pair *DCHS1-FAT4*. In this context, single-cell RNA-sequencing is applied to dissect aberrant gene expression programs at the level of specific cell types in the organoids as well as 2D cultures of neurons.

Chapter 1

Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development

The human neocortex underwent remarkable changes during evolution, most prominently an increase in size, which has been attributed to higher cognitive abilities in humans [223]. Thus, early developmental distinctions between species likely contribute to the differences that can be observed between humans, and our closest living relatives, the chimpanzees. In this study, cerebral organoids were adopted to study early development between human and chimpanzee using histology, live imaging (performed by collaborators in the group of Wieland Huttner) and single-cell RNA-Seq, to pinpoint cellular and molecular changes potentially contributing to evolutionary differences.

To decipher molecular differences between the species at cell type resolution, human and chimpanzee organoids were profiled using single-cell RNA-seq. Single-cell transcriptomic analyses of chimpanzee organoids showed clusters of cortical neural progenitor cells (NPCs) and neurons that could be aligned on a developmental pseudotime, very similar to single-cell transcriptomes from human organoids and fetal brain [114]. This demonstrated that the main lineage conversion from NPCs to neurons is conserved between the species and can be properly modeled in cerebral organoids.

Differential gene expression analysis between both species for major cell types showed that most cell type-specific genes were not differentially expressed between the species, confirming that most features are conserved. Nevertheless, a number of genes could be detected as differentially expressed between human and chimpanzee apical progenitors (APs) and neurons, respectively, with a similar number being more highly expressed in human cells and chimpanzee cells. Functional enrichment analysis for these differentially expressed genes identified membrane and intercellular signaling, and interesting candidate genes more highly expressed in humans were integrin beta 8 (*ITGB8*) in APs and insulin receptor (*INSR*) in neurons.

Moreover, inspecting the histological architecture in human and chimp organoids revealed very similar organization of germinal zones reminiscent of the developing fetal cortex with consistent numbers of cell types and cycling cells between both species. Measurements of the length of specific phases of the cell cycle using time-lapse imaging showed an intriguing elongation in metaphase in humans during mitosis, specifically in APs at early stages (day 30) of organoid development.

Overall, this work stresses the applicability of cerebral organoids as a useful in vitro model to perform comparative developmental studies, thus adding to previous 2D studies comparing species [59].

Chapter 2

Organoid single-cell genomic atlas uncovers human-specific features of brain development

Previous studies using cerebral organoids to study evolutionary development, have applied single-cell transcriptomics to decipher cell type-specific changes, however at lower throughput and without including many time points across differentiation to investigate stage-specific gene expression changes [20, 222]. In addition, current comparative studies of early development focused on gene expression, however investigating the underlying regulatory mechanisms still need to be explored.

In this study, we profiled thousands of cells from cerebral organoids generated from iPS and ES (embryonic stem) cells of human, chimpanzee and macaque using droplet-based microfluidic single-cell transcriptomics [224]. By analyzing the development of human and chimpanzee organoids from pluripotency to 4 months, we could reconstruct a trajectory where cells traverse different stages such as neuroectoderm, neuroepithelium and neural progenitor states until they diversify into different regional identities such as dorsal and ventral forebrain, midbrain and hindbrain. We extended the single-cell dataset of human organoids at 2 months to a total of 7 human lines and could observe different propensities of these lines for generating regional identities, similar to what has been described in previous research [113]. However, gene expression patterns of cells derived from the same organoid region showed high correlation, suggesting that gene expression programs were highly reproducible.

Reconstructing the same organoid developmental trajectory using open chromatin signatures from ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) data [197] revealed dynamic temporal changes in the regulatory landscape. To illuminate species-specific differences, pseudotime-based gene expression trajectories [225] of human and chimpanzee cells were aligned. This analysis identified cells in the chimpanzee organoids that did not map to a human equivalent, due to an increased maturation suggesting timing differences between the species [55, 59, 222]. Indeed, whereas in chimpanzee, upper layer neurons had already formed, only very few of those cells were observed in human organoids at the corresponding time points. Moreover, chimpanzee and macaque neurons appeared to be more mature compared to human neurons based on gene expression patterns related to neuron maturation.

By placing cells onto a pseudotime path [225] and aligning trajectories of 2-month old organoids generated from human, chimpanzee and macaque cells, cell type-specific patterns of differential gene expression (DE) were revealed and integrated with differential accessibility (DA) signals obtained from open chromatin profiling. This revealed cadherin 7 (*CDH7*) as well as a nearby putative regulatory region as interesting candidates for follow-up studies. The regulatory region overlapped with a human-accelerated region and single nucleotide changes fixed in humans, highlighting its potential evolutionary importance.

To dissect which species-specific gene expression changes persisted in the adult brain, single nuclei from frozen prefrontal cortex tissue from humans, chimpanzees/bonobo and macaques were investigated. The tissue was sliced from apical to basal positions and thus allowed for a spatial reconstruction of gene expression signatures. In general, glial cells, in particular astrocytes, showed the highest degree of gene expression changes and a lower conservation compared to neurons. Some gene expression differences, for example of the gene *COL6A1* were commonly found in organoids and the adult brain, thus emphasizing the relevance of organoids to study changes which possibly still impact later developmental phases. However, other genes were exclusively differentially expressed in either stage, thus stressing the power of complementary approaches using organoids and adult tissue to dissect brain differences.

Taken together, this study represents a comparative temporal atlas of gene expression and chromatin accessibility differences during early brain development of human, chimpanzee and macaque and links differences in the organoid to species differences in the adult brain, representing a valuable resource for the neuroscience community. It extends and complements recent investigations focusing on the human brain [116, 215, 219–221] and comparative single-cell studies using organoids with lower throughput [20, 222]. Future functional dissection of the candidates identified here has high potential to shed light on the multidimensional molecular basis of human-specific cognitive abilities.

Chapter 3

Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia

Brain development constitutes a finely orchestrated process in which aberrations lead to variable phenotypes and conditions. For example, erroneous behavior of neural progenitor cells or neurons can lead to various malformations of the cortex including anomalies in brain size, cortex folding and localization of neurons [17]. Periventricular neuronal heterotopia (PH) is a cortical malformation characterized by atypical locations of neurons along the ventricles in the brain, forming either clusters (nodular) or layers (laminar) of incorrectly placed neurons. However, the aberrant phenotype is limited to only a subgroup of neurons which remain at the ventricle and do not properly integrate into the cortex whereas a large fraction of the neurons behave properly. Patients presenting this pathology often suffer from epilepsy and intellectual disability [79, 226].

Investigations of genetic underpinnings of PH via sequencing of affected individuals highlighted *DCHS1* and *FAT4*, a cadherin receptor-ligand pair, as causal genes. Effects of mutations in *DCHS1* and *FAT4* were modeled in mice to shed light on the molecular mechanisms involved in disease etiology [79, 227]. However, mice did not fully recapitulate the phenotype observed in humans, potentially due to the differences in genetic background compared to humans and due to the fact that mice with their lissencephalic (smooth) cerebral cortex present a poor model for the gyrencephalic (convoluted) human cerebral cortex.

The availability of iPS cells reprogrammed from blood cells or fibroblasts opened the door to recapitulate disease mechanisms using patient-specific cells [80, 81] and new genome engineering tools such as CRIPSR/Cas9 allow for the introduction of patient-specific mutations in isogenic cell lines. To study PH, iPS cells from patient fibroblasts carrying the disease mutations were generated using reprogramming as well as isogenic *FAT4* and *DCHS1* knock-out and control stem cells using CRISPR/Cas9 genome editing. These cells were used to culture cerebral organoids, which have already proven useful for the study of diseases such as microcephaly and lissencephaly [92, 119].

Histological and cellular characterization of organoids produced from patient-derived and isogenic KO iPS lines (performed by our collaborators in the group of Silvia Cappello) indeed revealed phenotypes reminiscent of the PH condition, with neurons accumulating at the ventricles as well as disordered germinal zones. These observations were further corroborated by morphological changes of NPCs in the organoids and premature delamination from the apical membrane, thus leading to more differentiated neurons. Moreover, tracking the dynamic migration behavior of neurons in the organoids downregulated for *DCHS1* and *FAT4* and in 2D culture revealed a subgroup of neurons with an altered migratory phenotype.

To dissect the molecular mechanisms and gene regulatory networks potentially underlying the above described observations, single-cell RNA-sequencing of control and patient cerebral organoids was performed. Although the main characteristics of NPCs and neurons were similar between control and mutant organoids, mutant cells displayed a specific set of dysregulated genes, among them *RND3* whose role in neurogenesis and neuronal migration has been noted before [228]. Computational reconstruction of a developmental trajectory furthermore revealed a subgroup of neurons in the mutant organoids with upregulated gene expression patterns linking to axon guidance, neuronal migration and patterning including state-specific genes such as *ROBO3* and *CNTN2*. Contrarily, genes related to synapse formation and cytoskeleton associated genes were downregulated in this subpopulation. Interestingly, a similar subgroup enriched in mutant neurons was identified in neurons differentiated in a 2D culture system, corroborating the findings from histological and cellular analyses that showed a subpopulation of cells with changed migration behavior.

The underlying mechanisms as to why only a subset of neurons shows different migrational phenotypes even though harboring the same mutations, still need to be explored in upcoming studies. Nevertheless, combining cerebral organoids and single-cell RNA-seq holds great potential to decipher the mechanisms of neurodevelopmental diseases, such as heterotopia, and the knowledge gained from these studies could ultimately lead to targeting strategies for inaccurately organized neuronal circuits.

Conclusions and outlook

Single-cell sequencing and cerebral organoid technologies are influential technological innovations developed in recent years and have been used in this work to shed light on human-specific features of brain development resolved to distinct cell types and states. By comparing these differences to single-nucleus data generated from frozen adult tissue, it was possible to define differences that are specific to the developing brain as well as changes that still can be detected in the adult brain. Moreover, single cell RNA-seq was applied on patient-derived cerebral organoids to dissect the gene expression programs and mechanisms of neuronal heterotopia, a cortical malformation, whose phenotype was not fully recapitulated so far by using the mouse as a model system [79]. These studies thus emphasize the relevance of utilizing organoids to investigate comparative and aberrant brain development integrated with the benefit of single-cell sequencing to dissect gene expression programs at high resolution.

Studying primate organoids over a time course from pluripotent stem cells to 4 months has revealed timing differences in neuron maturation. Analyzing older organoids could inform about additional transcriptional or morphological differences in neuron maturation as well as in the timing of gliogenesis since astrocytes were only observed in the late stages of organoid development. Ongoing improvements in organoid cultivation protocols leading to enhanced maturation [94, 104, 128] could further aid in interrogating these later stages of development and would also help identifying pathogenic processes of more mature cell types when studying cortical malformations such as PH.

Recent studies have also measured electrical activity in organoids [118, 128] and comparing these features across different primates and during extended periods of cultivation could lead to insights about electrophysiological divergence during early brain development as well as potential timing differences in neuronal network establishment. Since cortical malformations often are accompanied by epileptical seizures, suggesting different electrical activity in neuron subpopulations, measuring electrical activity in patient-derived organoids represents a useful addition to examine how well disease conditions can be modeled in vitro [118, 128].

Assessing the transcriptome and chromatin accessibility in primate organoids has revealed an exciting set of cell type-specific candidate genes and regulatory regions likely playing a role in human-specific developmental divergence from chimpanzees, some of which retain into adulthood. Genes with expression changes could be further analyzed in organoids by using single-cell high throughput knock-out screens [177] to dissect the genes' function on gene regulatory networks. This would allow to isolate gene expression changes that have an effect during development and to narrow down candidates that could be investigated in more detail by using knock-outs introduced using CRISPR/Cas9 genome editing or overexpression of single genes. In addition, some genes with species-specific expression differences and accessibility overlap with interesting evolutionary signatures such as single nucleotide variants or human accelerated regions. Advances in single cell assays allow to test sets of regulatory elements in high troughput at the single cell level [229] in organoids. Moreover, genome editing would also allow for changing human-specific single nucleotide variants back to the chimpanzee state, which could be an interesting approach for *CDH7* that shows human-specific single nucleotide variants in its regulatory region. Moreover, it is higher expressed in human neurons compared to chimpanzee, thus potentially playing a role in neuron maturation.

Cells undergo complex developmental trajectories during development which could be disentangled by using lineage tracing methods with single cell readout [174, 175, 230, 231].

Comparing human and chimpanzee organoids using this method offers the possibility to reveal differences in the lineages that progenitor cells undergo to become neurons and glia cells. This could help to resolve the timing differences in neuronal development that were found in our study. Moreover, when comparing patient-derived and healthy organoids to study periventricular heterotopia, subsets of NPCs and neurons were identified by clustering and in silico lineage reconstruction. It is not clear how the cells that show altered behavior are related to the cells that have a normal phenotype despite carrying the same disease mutations. Lineage tracing would offer a useful approach to illuminate how these different cell types are related to one another, thereby gaining a better understanding of the mechanisms of the disease.

In addition, recent extensions of single-cell sequencing methods to interrogate additional modalities such as lncRNAs, microRNAs [232], chromosome conformation and methylation [207] would contribute to a multi-layered view on cell type specification and diversity [205, 206] both in the adult brain as well as in organoids from different species or patient-derived organoids. This would provide a more detailed view on the regulatory logic of cell type-specific gene expression differences in the context of evolution and disease.

Single-cell transcriptomics delivers a high resolution portrait of cell states and cell types, however, the spatial context is lost during dissociation of tissues. Consecutive sectioning of cubes from prefrontal cortex allowed for reconstructing spatial gene expression patterns of differentially expressed genes in this study. Validating these expression patterns in tissue using quantitative FISH (fluorescence in situ hybridisation) methods such as single molecule FISH would be a next question to address [233]. Furthermore, during the last years, spatial transcriptome profiling methods have been deployed that allow for the quantitative visualization of hundreds to thousands of transcripts in single cells [171, 173, 234]. This would allow to integrate morphological imformation and cell type-specific gene expression profiles.

Lastly, in the studies presented here, mainly human, chimpanzee, bonobo and macaque were studied to identify species differences at the transcriptome level. Thus, both for studying brain development in organoids as well as cellular diversity in the adult brain, including additional species, such as gorilla, orangutan or marmoset, and further individuals of the species studied here, would offer the possibility to dissect species-specific cell type and developmental timing characteristics in an even more finely resolved manner.

Zusammenfassung

Einleitung

Menschen leben in kompliziert organisierten Gesellschaften und haben Technologien entwickelt, die es ihnen ermöglichen mit unterschiedlichen Lebensräumen und Lebensbedingungen zurechtzukommen. Dies kann nur durch ein hohes Maß an Zusammenarbeit, Kommunikation über Sprache und komplexe Planung erreicht werden, die eine Anpassung und Wissensvermittlung über mehrere Generationen hinweg ermöglicht [1]. Obwohl unsere nächsten lebenden Verwandten, die Menschenaffen und insbesondere die Schimpansen, kognitive Fähigkeiten höherer Ordnung wie Werkzeuggebrauch und sogar Verhaltensweisen aufweisen, die mit Tradition und Kultur in Verbindung gebracht wurden [2, 3], sind moderne Menschen einzigartig was das Ausmaß ihrer technologischen Fortschritte und die Art und Weise, wie sie die Umwelt verändert haben, angeht. Diese kognitiven Unterschiede zwischen Mensch und Schimpanse haben eine Grundlage in Form, Funktion und Zellzusammensetzung im Gehirn, und diese Unterschiede haben wahrscheinlich einen Ursprung in frühen Entwicklungsphasen. Bislang ist jedoch wenig darüber bekannt, wie sich die Gehirnentwicklung zwischen Menschen und Menschenaffen unterscheidet. Zudem gestaltet es sich schwierig, die Mechanismen zu untersuchen, die den potenziellen Unterschieden zugrunde liegen, da es kein experimentelles System gibt, das die Entwicklung des Gehirns effektiv nachbildet.

Gehirnentwicklung und menschenspezifische Eigenschaften des Gehirns

Frühe Phasen der Gehirnentwicklung

Das Gehirn entwickelt sich aus dem Neuroektoderm, das die Neuralplatte bildet und sich anschließend zum Neuralrohr faltet. Das Neuralrohr segmentiert sich folgend in drei Vesikel, das Vorderhirn (Prosencephalon), das Mittelhirn (Mesencephalon) und das Hinterhirn (Rhombencephalon) einschließlich der Großhirnrinde, die aus dem Prosencephalon entsteht, und als Telencephalon bezeichnet wird. Das Neuralrohr besteht aus einer Einzelschicht von Zellen, die als Neuroepithel bezeichnet wird und Stammzellen beherbergt, die entlang einer apikobasalen Achse vom Lumen (apikale Seite) zur Außenseite (basale Seite) ausgerichtet sind. Neuroepithelzellen sind die Stammzellen des sich entwickelnden Gehirns, die sich durch Teilung zu radialen Gliazellen entwickeln. Diese wiederum weisen ein beschränktes Entwicklungspotenzial auf und differenzieren in verschiedene Vorläuferzelltypen.

Die verschiedenen kortikalen Vorläuferzellen unterscheiden sich in ihrem Proliferations-und Differenzierungspotenzial, der Position, an der die Mitose stattfindet, und dem Zellkontakt zur apikalen und basalen Membran [5]. Sie befinden sich dabei in unterschiedlichen Keimzonen des sich entwickelnden Kortex. Die apikalste Zone wird als ventrikuläre Zone (VZ) bezeichnet, gefolgt von der subventrikulären Zone (SVZ), der Intermediärzone/Mantelzone, der Rindenzone und der Marginalzone. Während der kortikalen Neurogenese entstehen aus den Vorläuferzellen exzitatorische glutamaterge Neuronen, die entlang der radialen Gliafasern von innen nach außen zu ihrer Position an der basalen Seite des Kortex wandern, wo sie schließlich sechs kortikale Schichten bilden, die durch unterschiedliche Zellzusammensetzungen gekennzeichnet sind [6–9]. Andere Zelltypen, wie z.B. die Mehrheit der GABAergen inhibierenden Neuronen (Interneuronen) im Kortex entwickeln sich aus anderen Keimzonen, den sogenannten ganglionischen Vorwölbungen, und wandern in den Neokortex, um sich in neuronale Schaltkreise zu integrieren [10–12].

Aus den radialen Gliazellen entwickeln sich später während der Entwicklung Gliazellen wie z.B. Astrozyten und myelinisierende Oligodendrozyten. Darüber hinaus werden neuronale Schaltkreise durch Synapsenbildung aufgebaut und verfeinert. Die Gehirnentwicklung ist bei der Geburt noch nicht abgeschlossen, viele Prozesse wie Hirnwachstum, die Ausbildung von Synapsen, die Gliogenese und Myelinisierung erstrecken sich über die fetale Entwicklung hinaus in die postnatale Phase. Die Myelinisierung zum Beispiel durchläuft sogar noch Veränderungen im erwachsenen Gehirn [4, 13]. Das erwachsene Gehirn beherbergt schließlich rund 86 Milliarden Neuronen [14], von denen 10-20 Milliarden Neuronen und 15-30 Milliarden Gliazellen in der Großhirnrinde zu finden sind [15], welche zusammen mehr als 80% der gesamten Hirnmasse ausmacht [16].

Probleme während der kortikalen Entwicklung führen zu verschiedenen Fehlbildungen, die z.B. durch Beeinträchtigungen des Vorläuferzellverhaltens, der neuronalen Migration und Schaltungsintegration verursacht werden können. Dadurch können Unterschiede in der Gehirngröße (Mikrozephalie, Makrozephalie), der Furchung des Kortex (z.B. Lissenzephalie) und der Fehllokalisierung von Neuronen (Heterotopie) beobachtet werden [17]. Viele dieser Erkrankungen führen zu epileptischen Anfällen und kognitiven Einschränkungen. Darüber hinaus wird vermutet, dass auch neuropsychiatrische Erkrankungen wie Autismus [18–21] und Schizophrenie [19] ihren Urprung in humanspezifischen Merkmalen der Gehirnentwicklung haben könnten.

Humanspezifische Merkmale im sich entwickelnden und adulten Gehirn

Eines der wichtigsten Merkmale, das das menschliche Gehirn von unseren nächsten lebenden Verwandten, den Menschenaffen, unterscheidet, ist die Vergrößerung des Neokortex. Folglich zeigt der Mensch eine etwa dreifache Zunahme der Hirnmasse im Vergleich zum Schimpansen [22]. Diese Entwicklung könnte auf Unterschiede im Zellzyklus, des zeitlichen Ablaufs der Entwicklung oder Veränderungen in den Eigenschaften neuronaler Stamm-und Vorläuferzellen zurückzuführen sein. So dauert beispielsweise die Neurogenese beim Menschen länger als bei Makaken mit 100 Tagen im Vergleich zu 60 Tagen [23]. Darüber hinaus wurde diese Vergrößerung zum Teil auf die SVZ zurückgeführt, die bei Primaten im Vergleich zu Mäusen dicker und weiterhin in einen inneren (ISVZ) und äußeren Abschnitt unterteilt ist, der durch eine innere Faserschicht getrennt wird [24]. Die SVZ beherbergt einen speziellen Zelltyp namens basale radiale Gliazellen (auch äußere radiale Gliazellen genannt), die eine erhöhte Proliferationsfähigkeit aufweisen und damit mehr Neuronen bilden, was möglicherweise zu einer erhöhten Hirngröße beiträgt [25, 26]. Basale radiale Gliazellen sind bei Mäusen [27] fast nicht vorhanden und weniger ausgeprägt bei Frettchen [26] und Makaken [28]. Zudem zeichnen sie sich durch spezifische Genexpressionsprofile [29], morphologische Variabilität [30], komplexe Entwicklungsbeziehungen [28] und die Aufrechterhaltung stammzellartiger und proliferativer Merkmale durch die Erzeugung einer eigenen proliferativen Nische [29, 31] aus. Abgesehen von der allgemeinen Vergrößerung des menschlichen Gehirns ist immer noch umstritten, ob einige Hirnregionen, wie z.B. der präfrontale Kortex, zuständig für höhere kognitive und soziale Fähigkeiten, sowie komplexe Planung [32], beim Menschen überproportional groß sind [33–38].

Neben der oben beschriebenen Größendifferenz des Gehirns wurden weitere morphologische, zelluläre, physiologische und molekulare Unterschiede im menschlichen Gehirn beschrieben. So zeigten beispielsweise magnetresonanztomographische Aufnahmen von Menschen-und Schimpansengehirnen eine erhöhte Plastizität in der kortikalen Organisation und Morphologie beim Menschen, was möglicherweise zu humanspezifischen Lernfähigkeiten beitragen könnte [39]. Im Frontalpol liegen die pyramidenförmigen Neuronen in Schicht III des Kortex beim Menschen weiter auseinander als bei Menschenaffen [40] und zeigen im Allgemeinen eine komplexere Morphologie [41]. Darüber hinaus zeichnet sich der Mensch durch eine Vergrößerung der oberen kortikalen Schichten (II+III) aus, die zu Veränderungn in der Signalgebung zwischen den kortikalen Regionen führen könnten [42].

Einige menschenspezifische Unterschiede im Gehirn könnten durch neue Zelltypen oder häufiger vorkommende Zelltypen verursacht werden. Van Economo-Neuronen, ein Pyramidenzelltyp, der bei vielen Säugetieren mit einem großen Gehirn zu finden ist, kommt beispielsweise beim Menschen häufiger vor als bei anderen Menschenaffen und ist beim Menschen durch einen größeren Zellkörper gekennzeichnet [43–45]. Durch Genexpressionsstudien in verschiedenen Regionen des Menschen-, Schimpansen-und Makakengehirns wurde zudem ein neuer Interneuron-Subtyp identifiziert, der Dopaminsynthesemarker exprimierte [46]. Dieser Zelltyp war im Vergleich zum Menschenaffen im menschlichen Striatum häufiger und im Neokortex ausschließlich im Menschen vorhanden [46]. Ein weiterer kürzlich identifizierter humaner Interneuron-Zelltyp, so genannte "Rosehip cells", die sich in der Schicht I des Kortex befinden, wurden auf der Grundlage von morphologischen und Genexpressionsanalysen entdeckt. Es ist jedoch noch nicht klar, ob dieser Zelltyp ausschließlich im Menschen vorkommt, da bisher nur Menschen und Mäuse verglichen wurden [47]. Auch die Morphologie und Funktion von

Astrozyten weisen Unterschiede auf, z.B. so genannte variköse Projektionsastrozyten in der Schicht V/VI, die bei Mäusen und niederen Primaten fehlen, aber beim Menschen und Schimpansen zu finden sind, kommen beim Schimpansen mit geringerer Größe und Komplexität vor [48]. Darüber hinaus hat der Mensch ein erhöhtes Glia-Neuron-Verhältnis über verschiedene Bereiche des frontalen Kortex entwickelt [49].

Vergleichende Untersuchungen zur Zusammensetzung von Metaboliten [50, 51] und Lipiden [52, 53] im Gehirn haben zudem Aufschluss über dynamische und artspezifische Unterschiede während der Gehirnentwicklung geliefert. So wurden beispielsweise Veränderungen im Glutamatstoffwechsel im präfrontalen Kortex des Menschen, aber nicht im Kleinhirn im Vergleich zu Schimpanse und Makak [50] gefunden. Weiterhin wurden grundsätzlich mehr Veränderungen der Metaboliten im präfrontalen Kortex beim Menschen im Vergleich zu Schimpansen, Makaken und Mäusen [51] festgestellt. Ähnliche Ergebnisse wurden bei der Untersuchung von Lipiden erzielt, bei denen mehr Veränderungen im Neokortex im Vergleich zum Kleinhirn festgestellt wurden [52].

Im Vergleich zu Schimpansen scheint die Entwicklung des menschlichen Gehirns verlangsamt zu sein, ein Phänomen, das auch als Neotenie bekannt ist. Es wird angenommen, dass die verlängerten Entwicklungsphasen mehr Flexibilität und längere Lernphasen ermöglichen könnten, um in komplexen sozialen Konstrukten zurechtzukommen [54]. Genexpressionsstudien haben gezeigt, dass bestimmte Gruppen von Genen ähnliche zeitliche Verschiebungen aufweisen, die Neotenie nahelegen, und vor allem an der Synapsenentwicklung beteiligt sind [55–57]. Darüber hinaus dauert die Eliminerung von Synapsen und die Myelinisierung beim Menschen länger [13, 58]. Untersuchungen in vitro legen nahe, dass ähnliche zeitliche Unterschiede auch bei der pränatalen Entwicklung eine Rolle spielen könnten [59], jedoch sind diese Prozesse bisher nicht gut verstanden.

Mögliche genetische Ursachen humanspezifischer Gehirnmerkmale

Die genetischen Grundlagen der humanspezifischen Merkmale des Gehirns sind noch nicht gut verstanden und nur einige Kandidaten wurden funktionell näher charakterisiert. So wurden beispielsweise humanspezifische Deletionen [60] und Genduplikationen [61] identifiziert, die mit evolutionären Gehirnunterschieden verbunden sind. Die humanspezifische segmentale Duplikation *ARHGAP11B* [62] erhöht nachweislich die Amplifikation der basalen radialen Gliazellen und führt zu einer Faltung des Kortex bei Mäusen, was auf eine potenzielle Rolle bei der Evolution der Hirngröße im Menschen hinweist. Ein weiteres bekanntes Beispiel ist *NOTCH2NL* [63, 64], das in radialen Gliazellen im fetalen menschlichen Gehirn exprimiert wird, an einer verlängerten radialen Gliaproliferation beteiligt ist und somit zu einer Vergrößerung des Gehirns geführt haben könnte. *SRGAP2C* eine weitere segmentale Duplikation induziert eine verzögerte Bildung sogenannter Dornfortsätze an Dendriten und führt zu einer höheren

Dornfortsatzdichte im Neokortex der Maus [65, 66]. Ein weiterer Kandidat, *GLUD2* [67], ein Gen, das beim Menschen und bei Menschenaffen vorhanden ist, könnte eine Rolle bei der Lipidbiosynthese während der postnatalen Gehirnentwicklung spielen.

Konservierte Sequenzen, die im Menschen verloren gegangen sind, können ebenfalls zu evolutionären Veränderungen beitragen. Eine dieser humanspezifisch konservierten Deletionen inaktiviert einen vermeintlichen Enhancer des Gens *GADD45G* und könnte aufgrund seiner Rolle bei der Gewebeproliferation somit für eine vergrößerte Gehirngröße verantwortlich sein. Im Allgemeinen scheinen menschliche konservierte Deletionen in der Nähe von Genen, die an der Neurogenese des Kortex beteiligt sind, überrepräsentiert zu sein. Häufig stehen sie im Zusammenhang mit Genen, die wahrscheinlich an einer erhöhten Proliferation beteiligt sind [60].

Das bisher herausragendste Beispiel auf der Ebene der Unterschiede in der Proteinkodierungssequenz ist *FOXP2*. Es ist durch zwei Aminosäureänderungen speziell im Menschen gekennzeichnet und führt zu Unterschieden in der Stimmbildung und beim Lernen in humanisierten Mäusen [68, 69]. Studien an Mausembryonen, die mit der humanen *FOXP2*-Variante elektroporiert wurden, deuteten darauf hin, dass *FOXP2* auch an der frühen Entwicklung beteiligt sein könnte, indem es das relative Verhältnis von radialen Gliazellen zu differenzierteren neuronalen Vorläufern ändert [70].

GPR56, ein Gen, das an der Proliferation von Vorläuferzellen und der kortikalen Furchung beteiligt ist, weist eine erhöhte Anzahl von Promotoren im Vergleich zu Mäusen auf. Die Deletion eines dieser regulatorischen Elemente führte zudem zu regional begrenzten Faltungsunterschieden im menschlichen Gehirn [71].

Human Accelerated Regions (HARs) sind genomische Segmente, die einen Überschuss an Sequenzänderungen in der menschlichen Linie aufweisen, aber in anderen Spezies konserviert sind [72]. Im Allgemeinen befinden sich viele dieser Regionen in nicht-kodierenden genomischen Regionen und wurden mit Enhancern, die in der Gehirnentwicklung aktiv sind, assoziiert [73]. Das Gen, welches die meisten HARs in seiner Nähe aufweist, heißt *NPAS3*. Es wurde bei Mäusen gezeigt, dass die menschliche Version einer dieser HARs, 2xHAR142, zu einem erweiterten Aktivitätsmuster führt, das auch das Vorderhirn im Vergleich zur Maus-und Schimpansenversion mit einschließt [74]. Ein weiteres interessantes Beispiel ist die *human accelerated region* 5 (*HARE5*), die die Genexpression von *FZD8* steuert. Bei Tests an Mäusen führt die menschliche Version zu einem schnelleren Zellzyklus bei neuronalen Vorläufern und einer Vergrößerung des Gehirns im Vergleich zur Schimpansensequenz [75]. HAR1 ist ein weiteres bekanntes Beispiel, das 18 Substitutionen auf der menschlichen Linie aufweist und sich in einem regulatorischen RNA-Gen, einer langen nichtcodierenden RNA (lncRNA), *HAR1F*, befindet. Die regulatorische RNA wird in der Großhirnrinde während der frühen Entwicklung in so genannten Cajal-Retzius-Zellen exprimiert, die eine wichtige Rolle beim Aufbau der laminaren Organisation der Großhirnrinde spielen. Somit ist *HAR1F* ein seltenes Beispiel für regulatorische RNAs, die eine Rolle bei der Evolution der menschlichen Gehirnentwicklung spielen könnten [72].

Die oben beschriebenen Studien haben somit geholfen, humanspezifische genetische Veränderungen zu identifizieren, die zu evolutionären Unterschieden in der Gehirnentwicklung führen. Die meisten der Versuche wurden jedoch bei Mäusen und nicht im humangenetischen Kontext durchgeführt. Daher wäre es wünschenswert, Modellsysteme zu verwenden, die die menschliche Biologie genauer nachbilden.

Fortgeschrittene in vitro Modellsysteme zur Untersuchung der Entwicklung

Bisherige Modelle und das Aufkommen induzierter Stammzellen

Es wird angenommen, dass viele der im menschlichen Gehirn zu beobachtenden evolutionären Veränderungen in der frühen Entwicklung verwurzelt sind, was die Bedeutung vergleichender Entwicklungsuntersuchungen unterstreicht. Die Erforschung der menschlichen Gehirnentwicklung, ihrer Fehlbildungen und Evolution beruht hauptsächlich darauf, dass entweder fetales Gewebe [76] oder Modellorganismen wie Mäuse [8], Ratten [6], Frettchen [77] oder Makaken [57] erforscht wurden. Die Verwendung von fetalem Gewebe ist jedoch mit ethischen Fragen und technischen Einschränkungen verbunden und Gewebe von Menschenaffen kann nicht untersucht werden. Darüber hinaus kann primäres Gewebe nicht ohne weiteres mit gentechnischen Methoden verändert werden, um die Wirkung ausgewählter Gene zu untersuchen. Transgene Modellorganismen wie Mäuse hingegen konnten die beim Menschen beobachteten Phänotypen oft nur teilweise widerspiegeln [78, 79], was höchstwahrscheinlich auf ihre phylogenetische Distanz zurückzuführen ist. Daher sind neue Modellsysteme erforderlich, die die menschliche Entwicklung genauer nachbilden, um vergleichende Entwicklung und Krankheitsprozesse genauer zu untersuchen.

Ein wichtiger Durchbruch auf dem Weg zu diesem Ziel war die Entdeckung, dass somatische Zellen aus vergleichsweise leicht zugänglichen Quellen wie Hautzellen [80] oder Blutzellen [81] durch die Verabreichung bestimmter Faktoren [80] zu sogenannten pluripotenten Zellen reprogrammiert werden können. Dadurch ist es möglich, Stammzellen von verschiedenen Primatenarten [82–84] und patientenspezifische Stammzellen zur Untersuchung von Krankheiten zu erzeugen [85]. Aufgrund ihrer Differenzierungskapazität können diese Zellen zu vielen verschiedenen Zelltypen entwickelt werden und bieten somit das Potenzial, verschiedene Gewebe von mehreren Individuen und Arten über mehrere Entwicklungsstufen hinweg zu untersuchen, um zeitlich aufgelöste Entwicklungprozesse zu entschlüsseln. Um humanspezifische Unterschiede in der Gehirnentwicklung zu untersuchen, wurden eine Reihe von Differenzierungsprotokollen in 2D entwickelt. Obwohl diese Systeme im Vergleich zum nativen Gewebekontext nur eine begrenzte [86–88] Komplexität aufweisen, haben sie sich beispielsweise als nützlich erwiesen, um zeitliche Unterschiede in der Entwicklung von Primaten zu identifizieren [59, 89].

Organoide als modernste 3D in vitro Modelle

In den letzten Jahren wurden so genannte Organoid-Technologien eingeführt, die es ermöglichen, frühe Entwicklungsstadien mit adulten, embryonalen oder induzierten pluripotenten Stammzellen nachzubilden, was das Feld der Stammzellforschung revolutionierte. Organoide sind 3D-Modelle, die sich zu organähnlichen Strukturen mit erhöhter Komplexität ausbilden, die gewebespezifische Zelltypen enthalten und Funktionen des Gewebes aufweisen [90]. Organoide können beispielsweise aus induzierten pluripotenten Stammzellen oder embryonalen Stammzellen des Menschen abgeleitet werden. Bis heute wurde eine Vielzahl verschiedener Gewebe nachgebildet, wie z.B. Darm [91], Gehirn [92–94], Leber [95], Niere [96], Lunge [97], Bauchspeicheldrüse [98], Netzhaut [99] und Innenohr [100]. Einer der Hauptvorteile, der sich durch die Verwendung von Organoiden aus Stammzellen gegenüber Primärgewebe ergibt, ist die Möglichkeit Gene, z.B. mittels CRISPR/Cas9, zu verändern [101, 102]. Dies ermöglicht die Untersuchung der Rolle bestimmter Gene in einem 3D-System, das Entwicklungsprozesse genauer nachahmt.

Arten von Gehirnorganoiden

Gehirnorganoide, die aus menschlichen embryonalen Stammzellen gewonnen werden, wurden 2008 von der Sasai-Gruppe [103] vorgestellt und später von Lancaster *et al.* [92] und Kadoshima *et al.* [93] weiter modifiziert. Das von Lancaster *et al.* eingeführte System produziert verschiedene Regionen des Gehirns in einem Organoid, was auch zu dem Begriff "zerebrale Organoide" für dieses System führte. Alternative Methoden wurden zudem dahingehend optimiert, bestimmte Regionen wie Vorderhirn [93, 104], Mittelhirn [104, 105] und Hypothalamus [104], Plexus choroideus und Hippocampus [106], sowie Hinterhirn [107] und Hypophyse [108] zu bilden. Zahlreiche Modifikationen und weitere Ergänzungen dieser Protokolle wurden eingeführt, darunter miniaturisierte Organoide, die in 3D-gedruckten Bioreaktoren gezüchtet werden, und so genannte Assembloide, um die Interaktion zwischen verschiedenen Gehirnregionen zu analysieren [109–112]. Die Untersuchung gewebespezifischer Merkmale [109–112], der Genexpression [20, 114–117], des Epigenoms [115, 116] und elektrischer Netzwerkaktivitäten [118] haben die Anwendbarkeit von Organoiden zur Untersuchung von Entwicklungsprozessen, die die frühe bis mittlere fetale Entwicklung beim Menschen nachbilden, zudem weiter bestätigt.

Anwendungsgebiete und Perspektiven

Zerebrale Organoide wurden bisher verwendet, um entwicklungsbedingte Störungen des Gehirns wie Mikrozephalie [92], kortikale Furchung [102, 119] oder neuronale Migrationsstörungen [109, 120], Down-Syndrom [121], sowie Autismus [122] zu studieren. So wurden beispielsweise Organoide zur Untersuchung des Miller-Dieker-Syndroms angewendet, einer Form der Lissenzephalie, die zu einer fehlenden Furchung des Kortex und einer reduzierten Hirngröße führt. Die Autoren fanden dabei Unterschiede in der Zellteilung äußerer radialer Gliazellen und einen vermehrten Zelltod in neuronalen Stammzellen [119]. Darüber hinaus wurden Organoide verwendet, um die Mechanismen der Infektion mit dem Zika-Virus (ZIKV) aufzuklären [104, 123, 124], ZIKV-Infektionshemmer zu testen [125], die Wirkung von Bisphenol A auf die Gehrinentwicklung zu untersuchen [104] und sogar um neurodegenerative Krankheiten wie die Alzheimer-Krankheit nachzubilden [126, 127]. Daher haben sich zerebrale Organoide bereits als sehr nützlich erwiesen, um verschiedene Fragen im Zusammenhang mit der Entwicklung des Gehirns und den Mechanismen von Krankheiten zu verstehen.

Die derzeitigen Nachteile von Organoiden stellen die begrenzte Reifungskapazität mit unvollständiger Schichtung kortikaler Regionen [104], sowie Größenbeschränkungen durch das Fehlen externer Signale und begrenzter Nährstoffversorgung dar. Darüber hinaus ist die Mehrheit der Protokolle durch Variabilität in der Zusammensetzung gekennzeichnet, inbesondere im Bezug auf das Vorkommen bestimmter Regionen, als auch durch chargenabhängige Unterschiede von Differenzierung zu Differenzierung [113, 128]. So zielen die jüngsten Bemühungen zur Verbesserung des Organoidsystems auf eine Erhöhung der Reproduzierbarkeit [117, 129], die Erzeugung von bestimmten Hirnregionen auf kontrollierte Art und Weise, sowie die Einführung derzeit fehlender Zelltypen wie Oligodendrozyten [130, 131] und Mikroglia [132] ab. Weiterhin soll die Reifung durch verlängerte Kultivierungszeiten [133], die Zugabe zusätzlicher Wachstumsfaktoren [104, 128] oder die Einführung von Blutgefäßen [134] verbessert werden, um die Anwendungsmöglichkeiten von Organoiden zu erweitern.

Die Erforschung der frühen Gehirnentwicklung bei Menschen und Menschenaffen stellt aufgrund ethischer und technischer Einschränkungen eine Herausforderung dar. Zerebrale Organoide, die aus pluripotenten Stammzellen generiert werden können, bieten jedoch die Möglichkeit, frühe Entwicklungsunterschiede zwischen diesen Arten in kontrollierten Zellkulturumgebungen zu modellieren. Einige Studien [20, 59] haben das Potenzial gezeigt, Unterschiede, die die frühe menschliche Gehirnentwicklung betreffen, zu detektieren, wie beispielsweise eine höhere Aktivierung des mTOR-Signalweges in menschlichen radialen Gliazellen im Vergleich zum Schimpansen [20].

Hochdurchsatz-Sequenzierungsstudien zur Gehirnentwicklung

Genexpressionsstudien

Mensch und Schimpanse haben 98,8% ihrer DNA-Sequenz gemein, jedoch sind trotz dieser hohen Ähnlichkeit eindeutige phänotypische Unterschiede erkennbar [135]. Es wurde daher vermutet, dass viele Merkmale, die den Menschen vom Schimpansen unterscheiden, auf nicht-kodierende regulatorische Sequenzunterschiede und Veränderungen in der Genexpression zurückzuführen sind und nicht auf Sequenzunterschiede in proteinkodierenden Genen [136, 137]. Damit stellen die Untersuchung genomweiter Muster der Genexpression und regulatorischer Signaturen vielversprechende Ansätze dar, um globale Veränderungen während der Gehirnentwicklung und im erwachsenen Gehirn zu analysieren, die den Menschen von den Menschenaffen unterscheiden.

Transkriptom-Sequenzierungsstudien des menschlichen Gehirns haben Veränderungen während der pränatalen [76, 138] und postnatalen Gehirnentwicklung [138] aufgedeckt und dynamische Genexpressionsänderungen mit einem Überschuss an Unterschieden in der pränatalen Entwicklung beschrieben [138]. In Studien, die postnatale Stadien der Gehirnentwicklung untersuchten, schienen die menschlichen Entwicklungsmuster vor allem im präfrontalen Kortex [139] größere Unterschiede aufzuweisen. Weiterhin war die Genexpression von Synapsen beim Menschen im Vergleich zu Schimpanse und Makake durch eine zeitliche Verschiebung zu späteren Entwicklungszeitpunkten gekennzeichnet [56]. Ebenso zeigte die Untersuchung des pränatalen und postnatalen Makakengehirns im Vergleich zum Menschen [57] eine längere Reifung im menschlichen Gehirn, die mit Genen verbunden war, die eine Rolle bei der Myelinisierung und Synaptogenese spielen [140].

Die Untersuchung schichtaufgelöster Genexpressionsmuster im präfrontalen Kortex von Mensch, Schimpanse und Makake lieferte Hinweise darauf, dass beim Menschen häufiger Unterschiede auftraten, die auf eine Reorganisation der Schichtexpressionsmuster hinweisen [141]. Die Erstellung von Genkoexpressionsnetzwerken, die eine integrierte Sicht auf Genbeziehungen liefern [142–145], zeigte, dass bestimmte Klassen von Transkriptionsfaktoren wie Krüppel assoziierte-Box Zink-Fingerproteine (KRAB ZNFs) bei Genexpressionunterschieden im präfrontalen Kortex zwischen Mensch und Schimpanse angereichert sind [143]. Zudem halfen diese Methoden humanspezifische Genexpressionsmodule im Frontalpol [144] zu identifizieren.

In einem größeren Kontext ist es jedoch noch nicht vollständig geklärt, ob das menschliche Gehirn einer beschleunigten Evolution unterlag, die auf einen Überschuss an Veränderungen in der menschlichen Linie hinweist, oder ob sich die Veränderungen linear über die evolutionäre Divergenzzeit anhäuften [146–150]. Darüber hinaus ist noch nicht bekannt, inwiefern bestimmte Zelltypen einer eventuell beschleunigten Evolution unterliegen könnten.

Unterschiede in der Genregulation

Eine Reihe von Studien untersuchte evolutionäre Veränderungen in der Genregulation wie z.B. DNA-Methylierung, Histonmodifikationen oder der Expression regulatorischer RNAs wie Mikro-RNAs und lncRNAs. Ein Vergleich der Histonmodifikationen zwischen Mensch, Makake und Maus im pränatalen Gehirn ergab zum Beispiel, dass menschenspezifische Regionen mit höherer Aktivität unter anderem mit der Zellproliferation und der extrazellulären Matrix assoziiert waren [151]. Eine ähnliche Studie, die im erwachsenen Gehirn von Mensch, Schimpanse und Makake durchgeführt wurde, beschrieb einen kleinen Satz von humanspezifischen aktiven Regionen, wobei 139 Regionen an Aktivität hinzugewannen und 45 Aktivität beim Menschen verloren hatten [152]. Vergleichbare Ergebnisse wurden erzielt, indem aktive Transkriptionsstartstellen in Neuronen des präfrontalen Kortex untersucht wurden, wobei 410 Loci hinzugewonnen wurden und 61 Loci ihre Aktivität speziell im Menschen verloren hatten [153]. Eine kürzlich durchgeführte Studie zeigte zudem die Verwendung von iPS-Zellen von Mensch und Schimpanse um diese zu Zellen der Neurralleiste zu differenzieren. Die Untersuchung von epigenomischen Mustern dieser Zellen könnte Unterschiede in Schädel-und Gesichtsform erklären, und unterstrich zudem die Anwendbarkeit von Stammzellmodellen zur Untersuchung evolutionärer Unterschiede [154]. Regulatorische RNAs wie microRNAs [46, 139, 155] oder lncRNAs [156] wurden bisher erst wenig im Zusammenhang mit humanspezifischen Gehirnunterschieden erforscht. Weiterhin könnte auch DNA-Methylierung eine Rolle bei Genexpressionsunterschieden spielen, wie durch den Vergleich der präfrontalen Hirnrinde von Mensch und Schimpanse untersucht wurde, die höhere Methylierung im Schimpansen zeigte [157].

Obwohl die oben genannten Studien wertvolle Ressourcen darstellen, um humanspezifische Genexpressionsunterschiede und ihre potenziell zugrunde liegende regulatorische Logik zu analysieren, basiert die Mehrheit auf Gesamtmessungen heterogener Gewebe. Auf diese Weise könnten einige der beobachteten Veränderungen auf unterschiedliche Zusammensetzungen im Gewebe zurückzuführen sein, sodass subtilere zelltypspezifische Veränderungen verdeckt werden.

Die Bandbreite von Einzelzellmethoden

Technologische Fortschritte der letzten Jahre haben es ermöglicht, die Genexpressionsprofile von Einzelzellen zu untersuchen [158]. Die Einzelzellanalyse ist aufgrund ihrer Fähigkeit, komplexe Gewebezusammensetzungen aufzulösen, seltene Zelltypen zu erkennen und orthologe Zelltypen zwischen verschiedenen Arten zu vergleichen [159] zu einer weit verbreiteten Methode geworden.

Die ersten Einzelzell-Transkriptommethoden basierten auf der manuellen Zellisolierung und der Generierung von Sequenzierbibliotheken in Reaktionsgefäßen mit geringem Durchsatz [158].

Eine Erhöhung des Durchsatzes wurde durch ventilbasierte Mikrofluidikgeräte, die auf die Trennung und Isolierung von Zellen ausgerichtet sind (Fluidigm C_1-System) [160–162], durch Zellsortierung und automatische Verarbeitung kleiner Volumina (MARS-Seq, massive parallele RNA-Einzelzellsequenzierung) [163], sowie die Verwendung von mikrostrukturierten Platten mit Nanowells zur Zellisolation erreicht [164]. Die Einführung von tröpfchenbasierten mikrofluidischen Techniken, die darauf beruhen, Tropfen als separate Reaktionsgefäße zu verwenden, hat jedoch den Durchsatz von Einzelzellstudien erheblich verbessert, so dass es möglich wurde, Tausende von Zellen in einem Experiment zu untersuchen [165, 166]. Zusätzliche Neuerungen, so genannte "Split-und-Pool"-Ansätze, haben die Anzahl der parallel untersuchbaren Zellen noch weiter erhöht und ermöglichen z.B. die Entdeckung seltener Zelltypen [167].

Diese Einzelzell-Transkriptomtechniken hatten somit bereits einen wesentlichen Einfluss auf die Untersuchung der Differenzierungsdynamik [161, 168], der Zellzustandsübergänge [169], der zellulären Zusammensetzung von Geweben [170–173], zellulärer Verwandschaftsbeziehungen [174, 175], sowie der Störung von zellulären Vorgängen [176, 177]. Jüngste gemeinsame Anstrengungen haben zudem damit begonnen mittels Einzelzell-Transkriptomik, Zellen aus allen wichtigen Organen von Mäusen [178] sowie von Menschen [179] zu untersuchen, um Referenzkarten von Zelltypen in gesunden [163, 169, 180–184] und pathologischen Zuständen [184–188] zu erstellen. Weiterhin wurden Zellreferenzen von häufig [189–192] und weniger häufig verwendeten Modellorganismen erstellt [168, 193]. Darüber hinaus hat sich die Einzelzellanalyse von der Untersuchung von Transkriptomen [194, 195] auf Genome, Epigenome [197, 198], Proteome, Chromosomen-Komformation [200, 201] und verschiedene Kombinationen selbiger [202–207] ausgeweitet, was eine integrierte Sicht auf Einzelzellzustände ermöglicht.

Einzelzellaufgelöste Untersuchungen des Gehirns

In den Neurowissenschaften hat sich die Einzelzell-Transkriptomanalyse zu einer Methode von unschätzbarem Wert entwickelt, um die Vielfalt der Gehirnzelltypen in der Maus [208–211] zu analysieren. Darüber hinaus wurden die Methoden erweitert, um die Einzelzellsequenzierung mit zusätzlichen Eigenschaften wie Zellmorphologie und elektrophysiologischen Eigenschaften zu kombinieren [212, 213], sowie Genexpressionsmuster aus gefrorenem und archiviertem Gewebe zu untersuchen [214–216]. Weiterhin wurde die Einzelzellanalyse angewendet, um die Mechanismen des sich entwickelnden [114, 217–220] und erwachsenen [215, 218, 221] menschlichen Gehirns zu beleuchten, einschließlich der Charakterisierung von 3D-Zellkulturmodellen sowohl von Menschen [20, 114, 117, 128], als auch Primaten [20, 222]. Zudem ergab ein Vergleich von Maus und Mensch gemeinsame und artspezifische Merkmale von Zellen im Gehirn [159].

Zielstellung der Dissertation und Überblick

Die Verfügbarkeit von Methoden zur Modellierung der menschlichen Gehirnentwicklung in kontrollierten Zellkultursystemem kombiniert mit Technologien zur Untersuchung von Geweben mittels Einzelzell-Transkriptomik, ermöglichen es, die Mechanismen zu verstehen, die der Evolution und Krankheitsenstehung des menschlichen Gehirns zugrunde liegen. In dieser Arbeit werden von Stammzellen abgeleitete zerebrale Organoide und Einzelzell-Transkriptomik genutzt, um humanspezifische Merkmale der Entwicklung des Kortex, sowie der Zelldiversität im adulten präfrontalen Kortex zu erforschen. Weiterhin werden Einzelzell-Transkriptomanalysen eingesetzt, um Zellpopulationen im Rahmen einer kortikalen Fehlbildung zu untersuchen.

Die ersten beiden Kapitel konzentrieren sich auf die Verwendung zerebraler Organoide, um die humanspezifische Gehirnentwicklung zu verstehen. In Kapitel 1 werden zerebrale Organoide, die aus induzierten pluripotenten Stammzellen von Mensch und Schimpanse erzeugt werden, verwendet, um frühe Entwicklungsänderungen im Menschen nachzubilden. Hierzu werden Einzelzell-Transkriptomanalysen verwendet, um zelltypspezifische Unterschiede zu entschlüsseln, die möglicherweise zu menschenspezifischen Zellmerkmalen beitragen.

Kapitel 2 erweitert das in Kapitel 1 beschriebene Projekt durch die Analyse von Einzelzell-Transkriptomen einer größeren Anzahl von Individuen, Arten, Zeitpunkten und Zellen. Dadurch werden dynamische Unterschiede untersucht, die während der frühen Gehirnentwicklung auftreten. Die beobachteten Genexpressionsänderungen werden weiterhin mit Unterschieden regulatorischer Signaturen mittels Chromatinzugänglichkeit integriert und ihre Konsistenz anhand von Analysen von Einzelzellkern-Transkriptomdaten von adultem Hirngewebe von Menschen, Schimpansen/Bonobos und Makaken getestet.

In Kapitel 3 werden zerebrale Organoide genutzt, um veränderte Zellzustände, die während einer kortikalen Fehlbildung, auftreten, zu verstehen. Zerebrale Organoide werden hierbei aus patientenabgeleiteten iPS-Zellen generiert und genutzt um periventrikuläre Heterotopie (PH) zu modellieren, die durch Mutationen im Cadherin-Rezeptor-Ligandenpaar *DCHS1-FAT4* verursacht werden. Einzelzell-RNA-Sequenzierung wird hierbei angewendent, um anormale Genexpressionsprogramme auf der Ebene spezifischer Zelltypen in Organoiden sowie 2D-Kulturen von Neuronen zu analysieren.

Kapitel 1

Unterschiede und Gemeinsamkeiten zwischen neuronalen Vorläuferzellen von Menschen und Schimpansen während der Entwicklung der Großhirnrinde

Der menschliche Neokortex veränderte sich während der Evolution stark, was vor allem in einer Größenzunahme zu beobachten ist, auf die höhere kognitive Fähigkeiten beim Menschen zurückzuführen sind [223]. Frühe Entwicklungsunterschiede zwischen den Arten tragen wahrscheinlich zu den Unterschieden bei, die zwischen Menschen und unseren nächsten lebenden Verwandten, den Schimpansen, beobachtet werden können. In dieser Studie wurden zerebrale Organoide eingesetzt, um die frühe Entwicklung zwischen Mensch und Schimpanse unter Verwendung histologischer Methoden, zeitaufgelöster Bildgebung (durchgeführt von Mitgliedern der Gruppe von Wieland Huttner) und der Einzelzell-RNA-Sequenzierung zu untersuchen, um zelluläre und molekulare Veränderungen zu identifizieren, die möglicherweise zu evolutionären Unterschieden beitragen.

Um molekulare Unterschiede zwischen den Spezies auf der Ebene von Zelltypen zu entschlüsseln, wurden menschliche und Schimpansen-Organoide mit Hilfe von Einzelzell-RNA-Sequenzierung untersucht. Einzelzell-Transkriptomanalysen von Schimpansen-Organoiden zeigten Gruppen kortikaler neuronaler Voläuferzellen und Neurone, die entlanger einer Pseudozeitache angeordnet werden konnten, in ähnlicher Weise wie bei Einzellzelltranskriptomen von menschlichen fetalen und Organoidzellen [114]. Dies deutete darauf hin, dass die Differenzierung von neuronalen Vorläuferzellen zu Neuronen zwischen den Spezies konserviert ist und in Organoiden nachgebildet werden kann.

Differentielle Genexpressionsanalyse zeigte, dass die meisten zelltypspezifischen Gene zwischen den Arten nicht unterschiedlich exprimiert waren, was wiederum bestätigt, dass die meisten Merkmale konserviert sind. Durch den direkten Vergleich von apikalen Vorläuferzellen und Neuronen zwischen Mensch und Schimpanse wurden für beide Gruppen ähnlich viele unterschiedlich hochexprimierte Gene detektiert. Funktionelle Kategorien für diese Gene umfassten Membran-und interzelluläre Signaltransduktion. Interessante Gene mit höherer Expression im Menschen waren Integrin beta 8 (*ITGB8*) in Vorläuferzellen und der Insulinrezeptor (*INSR*) in Neuronen.

Darüber hinaus ergab die Untersuchung der Histologie in menschlichen und Schimpansen-Organoiden eine sehr ähnliche Organisation von Keimzonen, die an den sich entwickelnden fetalen Kortex erinnerten, mit einer sehr ähnlichen Anzahl von Zelltypen und proliferierenden Zellen. Darüber hinaus zeigten zeitlich aufgelöste Messungen des Zellzyklus eine interessante

Verlängerung der Metaphase beim Menschen während der Mitose, insbesondere bei apikalen Vorläuferzellen in frühen Stadien (Tag 30) der Organoid-Entwicklung.

Diese Arbeit unterstreicht somit die Anwendbarkeit von zerebralen Organoiden als nützliches in vitro Modell zur Untersuchung der vergleichenden Gehirnentwicklung und ergänzt damit frühere 2D-Studien zum Vergleich von Arten [59].

Kapitel 2

Einzelzell-Genomik-Atlas von Organoiden zeigt humanspezifische Merkmale der Gehirnentwicklung auf

Frühere Studien, in denen zerebrale Organoide zur Untersuchung der evolutionären Entwicklung verwendet wurden, verwendeten Einzelzell-Transkriptomik zur Entschlüsselung zelltypspezifischer Veränderungen, jedoch mit geringerem Durchsatz und mit weniger Zeitpunkten, um stadienspezifische Genexpressionveränderungen zu untersuchen [20, 222]. Darüber hinaus fokussierten sich aktuelle vergleichende Studien zur frühen Entwicklung auf Messungen der Genexpression, sodass die zugrunde liegenden Regulationsmechanismen noch immer unklar sind.

In dieser Studie wurden Tausende von Zellen unter Verwendung tröpfchenbasierter Mikrofluidik aus zerebralen Organoiden untersucht, die aus iPS-Zellen und embryonalen Stammzellen von Menschen, Schimpansen und Makaken generiert wurden [224]. Durch die Analyse der Entwicklung von menschlichen und Schimpansenzellen vom Stadium der Pluripotenz bis zu 4 Monate alten Organoiden konnte eine Trajektorie rekonstruiert werden, in der Zellen verschiedene Stadien wie Neuroektoderm, Neuroepithelium und neuronale Vorläuferzustände durchlaufen, bis sie sich in verschiedene regionale Identitäten wie dorsales und ventrales Vorderhirn, Mittelhirn und Hinterhirn diversifizieren. Durch die Erweiterung des Datensatzes 2 Monate alter menschlicher Organoide auf 7 Zelllinien konnten unterschiedliche Tendenzen dieser Linien zur Ausbildung regionaler Identitäten aufgedeckt werden, was ein ähnliches Ergebnis zu früheren Untersuchungen darstellt [113]. Genexpressionsmuster von Zellen aus der gleichen Region wiesen eine hohe Korrelation auf, was darauf hindeutete, dass Genexpressionsprogramme innerhalb einer Region reproduzierbar waren.

Die Rekonstruktion der gleichen Trajektorie anhand von Signaturen zugänglichen Chromatins [197] zeigte zeitlich dynamische Veränderungen in regulatorischen Mustern. Um artspezifische Unterschiede aufzudecken, wurden pseudozeitbasierte Trajektorien [225] der Genexpression von Mensch und Schimpanse aneinander ausgerichtet. Diese Analyse identifizierte Zellen im Schimpansen, die aufgrund höherer Reife nicht zu einem menschlichen Äquivalent passten, was auf zeitliche Unterschiede in der Entwicklung zwischen den Arten hinweist [55, 59, 222]. Während sich bei Schimpansen bereits Neuronen der oberen Schichten gebildet hatten, wurden nur sehr wenige dieser Zellen bei menschlichen Organoiden zu den entsprechenden Zeitpunkten beobachtet. Darüber hinaus schienen Schimpansen-und Makakenneuronen basierend auf Genexpressionsmustern im Zusammenhang mit der Neuronenreifung reifer zu sein als menschliche Neuronen.

Durch die Anordnung von Zellen auf einem Pseudozeitpfad [225] und die Ausrichtung der Trajektorien von 2 Monate alten Organoiden, die aus menschlichen, Schimpansen-und Makakenzellen erzeugt wurden, konnten zelltypspezifische Muster der differentiellen Genexpression identifiziert und mit Signaturen unterschiedlicher Chromatinzugänglichkeit integriert werden. Dies ergab, dass Cadherin 7 (*CDH7*), sowie eine nahegelegene regulatorische Region, interessante Kandidaten für Folgeuntersuchungen darstellen. Die regulatorische Region überschnitt sich zudem mit einer im Menschen beschleunigt evolvierten Region und einzelnen, im Menschen fixierten Nukleotidänderungen, was deren evolutionäre Relevanz unterstreicht.

Um zu analysieren, welche artspezifischen Genexpressionsänderungen im erwachsenen Gehirn bestehen bleiben, wurden einzelne Zellkerne aus gefrorenem präfrontalem Kortexgewebe von Menschen, Schimpansen/Bonobos und Makaken untersucht. Das Gewebe wurde dabei von apikalen zu basalen Positionen geschnitten und ermöglichte so eine räumlich aufgelöste Rekonstruktion der Genexpressionssignaturen. Im Allgemeinen zeigten Gliazellen, insbesondere Astrozyten, den höchsten Grad an Genexpressionsänderungen und eine geringere Konservierung im Vergleich zu Neuronen. Einige Unterschiede in der Genexpression, wie beispielsweise das Gen *COL6A1*, wurden in Organoiden und im erwachsenen Gehirn gefunden, was die Relevanz von Organoiden für die Untersuchung von Veränderungen unterstreicht, die möglicherweise noch spätere Entwicklungsphasen beeinflussen. Andere Gene jedoch waren ausschließlich im sich entwickelnden oder adulten Gehirn unterschiedlich exprimiert, was die Vorteile komplementärer Ansätze unter Anwendung von Organoiden und adultem Gewebe zur Analyse von Gehirnunterschieden hervorhebt.

Zusammenfassend stellt diese Studie einen vergleichenden zeitlich ausgelösten Atlas von Genexpressions-und Chromatinzugänglichkeitsunterschieden während der frühen Gehirnentwicklung von Mensch, Schimpanse und Makake dar. Zudem verknüpft sie Unterschiede im Organoid mit Artenunterschieden im erwachsenen Gehirn und stellt somit eine wertvolle Ressource für andere Neurowissenschaftler dar. Die Studie erweitert und ergänzt neuere Untersuchungen mit Schwerpunkt auf dem menschlichen Gehirn [116, 215, 219–221] und vergleichende Einzelzellstudien mit geringem Durchsatz under Anwendung von Organoiden [20, 222]. Die zukünftige funktionelle Untersuchung der hier identifizierten Kandidaten hat hohes Potenzial, die multidimensionale molekulare Basis der humanspezifischen kognitiven Fähigkeiten näher zu beleuchten.

Kapitel 3

Verändertes neuronales Migrationsverhalten in menschlichen zerebralen Organoiden von Individuen mit neuronaler Heterotopie

Die Gehirnentwicklung ist ein fein orchestrierter Prozess, bei dem Abweichungen zu unterschiedlichen Phänotypen und Erkrankungen führen. So kann beispielsweise ein fehlerhaftes Verhalten neuronaler Vorläuferzellen oder Neuronen zu verschiedenen Missbildungen des Kortex führen, einschließlich Anomalien in der Hirngröße, Kortexfaltung und Fehllokalisierung von Neuronen [17]. Periventrikuläre neuronale Heterotopie (PH) ist eine kortikale Fehlbildung, die durch atypische Positionen von Neuronen entlang der Ventrikel im Gehirn gekennzeichnet ist und entweder Ansammlungen oder Schichten von falsch platzierten Neuronen bildet. Der anormale Phänotyp beschränkt sich jedoch nur auf eine Untergruppe von Neuronen, die an den Ventrikeln verbleiben und sich nicht richtig in den Kortex integrieren, die Mehrheit der Neuronen hingegen verhält sich normal. Patienten, die diese Pathologie aufweisen, leiden oft an Epilepsie und geistigen Einschränkungen [79, 226].

Untersuchungen der genetischen Grundlagen von PH durch Sequenzierung von betroffenen Individuen identifizierte *DCHS1* und *FAT4*, ein Cadherin-Rezeptor-Ligandenpaar als krankheitsauslösende Gene. Der Effekt dieser Gene wurde in Mäusen untersucht, um Licht in die molekularen Mechanismen der Krankheitenstehung zu bringen [79, 227]. Mäuse waren jedoch nicht in der Lage den beim Menschen beobachteten Phänotyp vollständig abzubilden, möglicherweise aufgrund der Unterschiede im genetischen Hintergrund im Vergleich zum Menschen und aufgrund der Tatsache, dass Mäuse mit ihrem glatten Kortex ein unzureichendes Modell für den menschlichen gefalteten Kortex darstellen.

Die Verfügbarkeit von iPS-Zellen, die aus Blutzellen oder Hautzellen umprogrammiert wurden, öffnete die Tür zur Rekapitulation von Krankheitsmechanismen unter Verwendung patientenspezifischer Zellen [80, 81]. Neue Werkzeuge zur Modifikation des Genoms wie CRIPSR/Cas9 ermöglichen zudem das Einbringen patientenspezifischer Mutationen in Zelllinien mit dem gleichen genetischen Hintergrund (isogen). Um PH zu untersuchen, wurden iPS-Zellen aus Patientenhautzellen, die die Krankheitsmutationen tragen mittels Reprogrammierung, sowie isogene *FAT4* und *DCHS1* Knock-Out und Kontrollstammzellen mit CRISPR/Cas9 Genomeditierung erzeugt. Diese Zellen wurden zur Kultivierung zerebraler Organoide verwendet, welche sich bereits bei der Untersuchung von Entwicklungserkrankungen des Gehirns wie Mikrozephalie und Lissenzephalie als nützlich erwiesen haben [92, 119].

Die histologische und zelluläre Charakterisierung von Organoiden, die aus patientenabgeleiteten und isogenen Knock-Out-iPS-Linien hergestellt wurden (durchgeführt von Kooperationspartnern in der Gruppe von Silvia Cappello), zeigte tatsächlich Phänotypen, die an die Erkrankung erinnern. Beispielsweise sammelten sich Neuronen an den Ventrikeln an. Außerdem wiesen die Keimzonen eine gestörte Organisation auf. Diese Beobachtungen wurden durch morphologische Veränderungen neuronaler Vorläuferzellen in den Organoiden und vorzeitiges Ablösen von der apikalen Membran, was zu differenzierteren Neuronen führte, weiter bestätigt. Darüber hinaus wurde das dynamische Migrationsverhalten von Neuronen in Organoiden verfolgt, bei denen *DCHS1* und *FAT4* heruntergeregelt wurde, sowe in 2D-Kulturen von patientenabgeleiteten Zellen. Dabei wurde eine Untergruppe von Neuronen identifiziert, die einen veränderten migratorischen Phänotyp aufwies.

Um die molekularen Mechanismen und Genregulationsnetzwerke, die den oben beschriebenen Beobachtungen zugrunde liegen könnten, zu analysieren, wurden Einzelzell-RNA-Sequenzierungen von zerebralen Organoiden von Patienten und Kontrollen durchgeführt. Obwohl die Hauptmerkmale neuronaler Vorläuferzellen und Neuronen von Kontroll-und Patienten-Organoiden ähnlich waren, zeigten mutierte Zellen einen spezifischen Satz dysregulierter Gene, darunter *RND3*, dessen Rolle bei der Neurogenese und neuronalen Migration bereits zuvor [228] beschrieben wurde. Die computergestützte Rekonstruktion des Differenzierungsverlaufs zeigte zudem eine Untergruppe von Neuronen in den mutierten Organoiden mit hochregulierten Genexpressionsmustern. Diese waren mit der Orientierung von Axonen, neuronaler Migration und Musterbildung verknüft. Zudem exprimierten sie Gene wie *ROBO3* und *CNTN2*, die spezifisch für diese Untergruppe waren. Im Gegensatz dazu waren in dieser Subpopulation beispielsweise Gene herunterreguliert, die mit Synapsenbildung und Zytoskelett assoziiert waren. Interessanterweise wurde eine ähnliche Untergruppe, die viele mutierte Neurone aufweist, in Neuronen gefunden, die in einem 2D-Kultursystem differenziert wurden. Dies bestätigte Ergebnisse der histologischen und zellulären Analysen, die eine Untergruppe von Zellen mit verändertem Migrationsverhalten zeigten.

Warum nur eine Untergruppe von Neuronen unterschiedliche Migrationsphänotypen zeigt, obwohl sie die gleichen Mutationen aufweiset, muss noch in weiterführenden Studien untersucht werden. Dennoch birgt die Kombination von zerebralen Organoiden und Einzelzell-RNA-Sequenzierung großes Potenzial, die Mechanismen von Entwicklungsstörungen wie Heterotopie zu entschlüsseln. Die Erkenntnisse aus diesen Studien könnten letztendlich zu zielgerichteten Behandlungsstrategien für fehlorganisierte neuronale Schaltkreise führen.

Schlussfolgerungen und Ausblick

Einzelzell-Sequenzierungstechnologien und zerebrale Organoide stellen einflussreiche technologische Innovationen der letzten Jahre dar und wurden in dieser Arbeit verwendet, um Aufschluss über humanspezifische zelltypaufgelöste Unterschiede zu geben. Durch den Vergleich dieser Unterschiede mit Einzelzellkerndaten aus gefrorenem adultem Gewebe gelang es, Merkmale zu definieren, die spezifisch für das sich entwickelnde Gehirn sind, sowie Veränderungen, die noch im erwachsenen Gehirn fortbestehen. Darüber hinaus wurde die Einzelzell-RNA-Sequenzierung bei patientenabgeleiteten zerebralen Organoiden angewendet, um die Genexpressionsprogramme der neuronalen Heterotopie, einer kortikalen Fehlbildung, zu analysieren. Deren Phänotyp war bisher nicht vollständig im Mausmodell nachstellbar [79]. Diese Studien unterstreichen somit die Relevanz der Verwendung von Organoiden zur Untersuchung der vergleichenden und anormalen Gehirnentwicklung, integriert mit dem Nutzen der Einzelzellsequenzierung zur Analyse von Genexpressionsprogrammen mit hoher Auflösung.

Die Untersuchung von Primatenorganoiden über einen Zeitraum von pluripotenten Stammzellen bis zu 4 Monaten hat zeitliche Unterschiede in der Neuronenreifung aufgezeigt. Die Analyse älterer Organoide könnte über zusätzliche Unterschiede in der Neuronenreifung sowie dem Zeitpunkt der Gliogenese aufklären, da Astrozyten erst im Spätstadium der Organoidentwicklung beobachtet wurden. Ständige Optimierungen der Organoid-Kultivierungsprotokolle, die zu einer verbesserten Reifung führen, könnten bei der Untersuchung dieser späteren Entwicklungsstadien helfen und wären zudem nützlich, um pathogene Prozesse reiferer Zelltypen beim Studium kortikaler Fehlbildungen zu identifizieren.

Neuere Studien haben zudem elektrische Aktivität in Organoiden gemessen [118, 128] und der Vergleich dieser Merkmale zwischen verschiedenen Primaten und während längerer Kultivierungsperioden könnte zu Erkenntnissen über elektrophysiologische Unterschiede während der frühen Gehirnentwicklung, sowie über mögliche zeitliche Unterschiede im Aufbau neuronaler Netzwerke informieren. Da kortikale Fehlbildungen oft von epileptischen Anfällen begleitet werden, was auf eine veränderte elektrische Aktivität in neuronalen Subpopulationen hindeutet, stellt die Messung der elektrischen Aktivität in patientenabgeleiteten Organoiden eine sinnvolle Ergänzung dar, um zu untersuchen, wie gut sich Krankheitsbilder in vitro nachbilden lassen [118, 128].

Die Untersuchung der Transkriptome, sowie der Chromatinzugänglichkeit in Primatenorganoiden hat eine Reihe von zelltypspezifischen Kandidatengenen und regulatorischen Regionen ergeben, die wahrscheinlich eine Rolle bei der menschenspezifischen Entwicklungsdivergenz im Vergleich zum Schimpansen spielen und einige dieser Unterschiede bestehen sogar bis ins Erwachsenenalter. Gene mit Expressionsänderungen könnten in Organoiden weiter analysiert werden, indem Einzelzell-Knock-Out-Screens mit hohem Durchsatz verwendet werden, um die Funktion der Gene in genregulatorischen Netzwerken zu analysieren [177]. Dies würde es ermöglichen, Genexpressionsänderungen zu identifizieren, die eine Rolle in der Entwicklung spielen und Kandidaten einzugrenzen, die durch Knock-Outs mittels CRISPR/Cas9 Genomveränderung, oder Überexpression einzelner Gene, genauer untersucht werden könnten. Darüber hinaus überschneiden sich einige Gene mit artspezifischen Expressionsunterschieden und regulatorischen Unterschieden mit interessanten evolutionären Signaturen wie Einzelnukleotidvarianten oder beschleunigt evolvierenden Regionen im Menschen. Fortschritte bei Einzelzell-Assays ermöglichen es, Gruppen von regulatorischen Elementen im hohen Durchsatz auf Einzelzellebene [229] in Organoiden zu testen. Darüber hinaus würde die Genomeditierung auch die Änderung humanspezifischer Einzelnukleotidvarianten zurück in den Schimpansenzustand ermöglichen, was ein interessanter Ansatz für *CDH7* sein könnte, das humanspezifische Einzelnukleotidvarianten in seiner regulatorischen Region zeigt. Darüber hinaus ist es in menschlichen Neuronen im Vergleich zu Schimpansen höher exprimiert und spielt somit möglicherweise eine Rolle bei den beobachteten Unterschieden in der Neuronenreifung.

Zellen durchlaufen während der Entwicklung komplexe Entwicklungstrajektorien, die durch die Verwendung von Methoden zur Abstammung dieser Zellen, sogenanntes *lineage tracing* mit Einzelzellauflösung [174, 175, 230, 231] untersucht werden könnten. Der Vergleich von menschlichen und Schimpansen-Organoiden mit dieser Methode bietet z.B. die Möglichkeit, Unterschiede in den Abstammungslinien aufzudecken, die Vorläuferzellen durchlaufen, um zu Neuronen und Gliazellen zu werden. Dies ist besonders interessant, da zeitliche Unterschiede in der neuronalen Entwicklung zwischen Mensch und Schimpanse in Organoiden gefunden wurden. Darüber hinaus wurden beim Vergleich von patientenabgeleiteten und gesunden Organoiden zur Untersuchung der periventrikulären Heterotopie veränderte Gruppen neuronaler Vorläuferzellen und Neuronen identifiziert, jedoch durch computergestützte Rekonstruktion der Differenzierung der Zellen. *Lineage tracing* würde es ermölichen die Verwandschaftsbeziehungen der Zellen auf direkterem Wege zu rekonstruieren. Dadurch könnte der Zusammenhang von Zellen, die sich anders verhalten und denen, die einen normalen Phänotyp trotz gleicher genetischer Ausstattung aufzeigen, näher beleuchtet werden, um die Mechanismen der Erkrankung besser zu verstehen.

Weiterhin können neue Erweiterungen von Einzelzellsequenzierungsmethoden zur Analyse zusätzlicher Modalitäten wie lncRNAs, microRNAs [232], Chromosomenkonformation und Methylierung [207] zu einem vielschichtigen Blick auf Zelltypeigenschaften und deren Vielfalt [205, 206] sowohl im erwachsenen Gehirn als auch in Organoiden verschiedener Spezies oder von Patienten abgeleiteten Organoiden beitragen. Dies würde eine detaillierte Sicht auf die regulatorische Mechanismen von zelltypspezifischen Unterschieden im Kontext von Evolution und Erkrankungen liefern.

Die Einzelzell-Transkriptomik liefert eine hochaufgelöste Beschreibung von Zellzuständen und Zelltypen, allerdings geht der räumliche Zusammenhang bei der Dissoziation von Gewebe für diese Experimente verloren. Die Validierung dieser Expressionsmuster im Gewebe unter

Verwendung quantitativer räumlich aufgelöster FISH-Methoden (Fluoreszenz in situ Hybridisierung) wie z.B. mittels Einzelmolekül-FISH stellt eine interessante Erweiterung dar [233]. Darüber hinaus wurden in den letzten Jahren räumlich aufgelöste Transkriptommethoden eingesetzt, die eine quantitative Visualisierung von Hunderten bis Tausenden von Transkripten in einzelnen Zellen ermöglichen [171, 173, 234]. Dies würde die Integration von Morphologie, sowie von zelltypspezifischen Hochdurchsatz-Genexpressionsprofilen ermöglichen.

In den hier vorgestellten Studien wurden hauptsächlich Menschen, Schimpansen, Bonobos und Makaken studiert, um Artenunterschiede auf der Transkriptomebene zu identifizieren. Die Untersuchung zusätzlicher Arten wie Gorilla, Orang-Utan oder Krallenäffchen, und weiterer Individuen der hier untersuchten Arten würde die Möglichkeit bieten, artspezifische Zelltypunterschiede, sowie Änderungen im zeitlichen Entwicklungsverlauf noch genauer zu analysieren.

Chapter 1

Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development

Felipe Mora-Bermúdez, Farhath Badsha, Sabina Kanton, J. Gray Camp, Benjamin Vernot, Kathrin Köhler, Birger Voigt, Keisuke Okita, Tomislav Maricic, Zhisong He, Robert Lachmann, Svante Pääbo, Barbara Treutlein, Wieland B. Huttner

eLife, Vol. 26, Issue 5 (2016)

Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development

Felipe Mora-Bermúdez[1]†, Farhath Badsha[1]†, Sabina Kanton[2]†, J Gray Camp[2]†, Benjamin Vernot[2], Kathrin Köhler[2], Birger Voigt[3], Keisuke Okita[4], Tomislav Maricic[2], Zhisong He[5], Robert Lachmann[6], Svante Pääbo[2]*, Barbara Treutlein[1,2]*, Wieland B Huttner[1]*

Affiliations:

[1]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

[2]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

[3]Institute of Laboratory Animals, Graduate School of Medicine, Kyoto University, Kyoto, Japan

[4]Department of Reprogramming Science, Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan

[5]CAS-MPG Partner Institute for Computational Biology, Shanghai, China

[6]Universitätsklinikum Carl Gustav Carus, Klinik und Poliklinik für Frauenheilkunde und Geburtshilfe, Technische Universität Dresden, Dresden, Germany

Correspondence to: paabo@ eva.mpg.de (SP),
barbara_treutlein@eva.mpg.de (BT),
huttner@mpi-cbg.de (WBH)

†These authors contributed equally to this work

Abstract

Human neocortex expansion likely contributed to the remarkable cognitive abilities of humans. This expansion is thought to primarily reflect differences in proliferation versus differentiation of neural progenitors during cortical development. Here, we have searched for such differences by analysing cerebral organoids from human and chimpanzees using immunohistofluorescence, live imaging, and single-cell transcriptomics. We find that the cytoarchitecture, cell type composition, and neurogenic gene expression programs of humans and chimpanzees are remarkably similar. Notably, however, live imaging of apical progenitor mitosis uncovered a lengthening of prometaphase-metaphase in humans compared to chimpanzees that is specific to proliferating progenitors and not observed in non-neural cells. Consistent with this, the small set of genes more highly expressed in human apical progenitors points to increased proliferative capacity, and the proportion of neurogenic basal progenitors is lower in humans. These subtle differences in cortical progenitors between humans and chimpanzees may have consequences for human neocortex evolution.

Introduction

The expansion of the neocortex during primate evolution is thought to contribute to the higher cognitive capacity of humans compared to our closest living relatives, the great apes, and notably the chimpanzees [23, 235, 236]. Neocortex expansion in humans relative to chimpanzees involves an increase in the number of cortical neurons generated during fetal development [237–240]. This reflects primarily a greater and prolonged proliferative capacity of human neural stem and progenitor cells (NSPCs) within the germinal zones of the developing neocortex [241]. Unravelling differences between human and chimpanzee NSPC behaviour is therefore a key issue, yet very little is known about such differences.

The neocortex develops from two principal germinal zones, the ventricular zone (VZ) and the subventricular zone (SVZ) [242]. In primates developing a folded (gyrencephalic) neocortex, and notably in humans, an inner SVZ (iSVZ) and an outer (oSVZ) can be distinguished [24, 243]. Correspondingly, the VZ and SVZ harbour the cell bodies of two principal classes of NSPCs, called apical progenitors (APs) and basal progenitors (BPs), respectively, each of which comprise several distinct NSPC types [5, 237, 240, 244]. APs (neuroepithelial cells, apical radial glia, and apical intermediate progenitors) divide at the ventricular surface, keep ventricular contact and exhibit apical cell polarity, whereas BPs (basal (or outer) radial glia and basal intermediate progenitors) lack this contact and type of cell polarity [5].

Studies dissecting the switch between NSPC proliferation and differentiation have demonstrated that a central aspect of the cell division process, the orientation of the mitotic spindle, has a pivotal role, particularly in the case of APs [245–248]. The orientation of the spindle relative to the apical-basal axis of cell polarity in mitotic apical radial glia, the major cortical neural stem cells [244, 249], can determine whether their division is symmetric or asymmetric, and whether it is proliferative or neurogenic, with regard to their progeny [245–248]. Comparing spindle orientation in mitotic APs may therefore provide insight into the cell biological basis underlying the differences between humans and chimpanzees in NSPC proliferation versus differentiation during neocortex development.

Protocols to generate structured cerebral tissue (cerebral organoids) from pluripotent stem cells in vitro constitute a major advance for studying neocortex development, in particular with regard to humans and non-human primates where fetal brain tissue is hard or impossible to obtain and manipulate [92, 93, 104, 122, 250]. Human cerebral organoids form a variety of tissues that resemble specific brain regions, including the cerebral cortex, ventral forebrain, midbrain-hindbrain boundary, hippocampus, and retina. Moreover, their cerebral cortex-like regions exhibit distinct germinal zones, that is, a VZ containing APs and an SVZ containing BPs, as well as basal-most neuronal layers. Cerebral organoid APs include apical radial glia-like NSPCs that contact a ventricle-like lumen, express radial glia marker genes, undergo interkinetic nuclear migration, and divide at the apical surface, similar to their in vivo counterparts, and cerebral organoid BPs comprise both basal radial glia-like and basal intermediate progenitor-like NSPCs [92]. Finally, we have previously shown by single-cell RNA sequencing that the gene expression programs controlling neocortex development in human cerebral organoids are remarkably similar to those in the developing fetal tissue [114]. Together, these findings suggest that cerebral organoids constitute a valid system to explore potential differences in NSPC proliferation versus differentiation between humans and chimpanzees [59], in particular with regard to spindle orientation in mitotic APs.

Here, we have generated cerebral organoids from chimpanzee-derived induced pluripotent stem cells (iPSCs), and used single-cell transcriptomics, immunohistofluorescence and live imaging to compare relevant features of chimpanzee NSPCs to human NSPCs in cerebral organoids and fetal neocortex. While most NSPC characteristics are found to be similar, we show that the prometaphase-metaphase in mitotic APs is longer in humans than in chimpanzees, indicating that a fundamental difference exists in the regulation of mitosis during neocortex development between the two species. Our data also provide a resource for further studies on human and chimpanzee differences in cortical development, and demonstrate the usability of cerebral organoids as a means to be able to perform such studies.

Results

Chimpanzee cerebral organoids recapitulate cortex development

We generated cerebral organoids from iPSCs derived from chimpanzee fibroblasts and lymphocytes (Figure 1A left, Figure 1—figure supplement 1). These chimpanzee cerebral organoids formed complex tissue structures that resembled the developing primate brain (Figure 1A right), as reported previously for human cerebral organoids [92]. Similar to human iPSC-derived cerebral organoids ([114], Figure 1B,C right), within the chimpanzee organoids grown for 52 days (D52), we observed cortex-like regions (Figure 1A right) with PAX6-positive APs (such as radial glia) residing predominantly in the apical-most zone facing a ventricular lumen (Figure 1B left), similar to the ventricular zone (VZ) of developing primate neocortex at an early-mid stage of neurogenesis. Consistent with this, cells immunoreactive for the deep-layer neuron marker CTIP2 were observed in the basal region of the developing cortical wall (Figure 1B left), corresponding to an early cortical plate. TBR2 (also known as EOMES) positive BPs (presumably mostly basal intermediate progenitors) were concentrated in a zone between the PAX6+ progenitors and the CTIP2+ neurons, corresponding to the subventricular zone (SVZ). In the context of the time-lapse live imaging of apical mitoses described below, we observed apically directed nuclear migration prior to, and basally directed nuclear migration after, mitosis, consistent with the existence of interkinetic nuclear migration. Our results suggest that chimpanzee cerebral organoids recapitulate important aspects of fetal chimpanzee brain development and allow comparisons with cerebral cortex development in human cerebral organoids and fetal neocortex.

We next compared the proportion of various NSPC types, as revealed by expression of PAX6 and/or TBR2, and neurons at a very early (D28) and a mid-neurogenic (D52-D54) stage between human and chimpanzee cerebral organoid cortices (Figure 2). In both species, we observed a decrease in PAX6+TBR2– apically located NSPCs (presumably proliferating APs) from D28 to D52, concomitant with an increase in PAX6+TBR2+ and PAX6–TBR2+ basally located NSPCs (presumably neurogenic BPs) (Figure 2A,B). Notably, whereas no significant differences were observed at D28, at D52-D54, the proportion of PAX6+TBR2+ NSPCs in the chimpanzee organoids was nearly twice that in the human organoids, and the proportion of PAX6+TBR2– NSPCs was correspondingly lower, whereas no significant difference between human and chimpanzee was observed for PAX6–TBR2+ NSPCs (Figure 2B). In line with what would be expected with regard to neuron production, the proportion of PAX6–TBR2– cells, located in the basal-most zones of the developing cortical wall, was very low at D28 but increased by D52-D54 to about a third of the total cells for both, human and chimpanzee cerebral organoids (Figure 2B). Immunostaining for CTIP2 corroborated the neuronal identity of these cells (data not shown).

Consistent with the observation that the total proportion of NSPCs relative to neurons was virtu- ally identical in human and chimpanzee organoids (Figure 2B), the abundance of cycling cells, as revealed by KI67 immunostaining, was essentially similar (Figure 2C,D). We conclude that at the two stages studied, there are – with the exception of the PAX6+TBR2+ NSPCs – no major differences between human and chimpanzee cerebral cortex developing in organoid culture with regard to the types of NSPCs and their abundance, or neuron output.

Cell composition and lineage relationships in chimpanzee cerebral organoids

To survey the cellular composition and cell type-specific transcriptomes of the chimpanzee organoids, we analysed 344 single cell transcriptomes from 7 organoids ranging in age from 45 to 80 days (Figure 1D, Figure 1—source data 1). We combined all transcriptomes and identified the genes most informative for defining cell populations by principal component analysis (PCA) (Figure 1—source data 2). Using these genes, we used tSNE analysis to cluster cells into transcriptionally distinct groups representing cerebral cortex, hindbrain, ventral midbrain and peripheral mesenchyme (Figure 1—figure supplement 2). These groups are similar to those identified in human cerebral organoids [114]. We identified 178 cortex-like cells based on strong expression of canonical NSPC and neuron marker genes (i.e., NSPCs and neurons: FOXG1, NFIA, NFIB; NSPCs: PAX6, SOX2, GLI3; neurons: NEUROD6) and the lack of expression of the ventral and medial telencephalic markers OTX2 and RSPO2 (Figure 1—figure supplement 2).

We sub-classified the 178 cerebral cortex-like cells based on the correlation between their transcriptomes and the bulk transcriptomes of laser-capture microdissected VZ, iSVZ, oSVZ, and cortical plate of fetal human neocortex (GSE38805, [31]). We found that groups of cells correlated best with one of the four zones, suggesting that the range of cell types present in the human fetal and organoid cerebral cortex are represented in our chimpanzee data (Figure 1E). Consistent with this, each chimpanzee cell represents a cell state on a continuum from NPSCs to neurons based on gene expression signatures extracted from fetal human cerebral cortex transcriptomes (Figure 1F, Figure 1—figure supplement 3) [114].

We next classified the chimpanzee cerebral cortex cells by determining the fetal cell type with which each cell most strongly correlates, resulting in 73 APs, 25 BPs, and 80 neurons. Analysis of known cell type markers revealed expression patterns consistent with what has been observed in human organoid and fetal cerebral cortex (Figure 1G) [114]. Though this classification is convenient to describe the cell types present in the chimpanzee organoid, we note that many of the cells can be described as intermediates between APs, BPs, and different stages of neuron maturation. We inferred lineage relationships among the chimpanzee cerebral cortex in an adjacency network based on pairwise correlations between cells (Figure 1H), revealing a structured topology where VZ-APs connect to cortical plate neurons through SVZ-

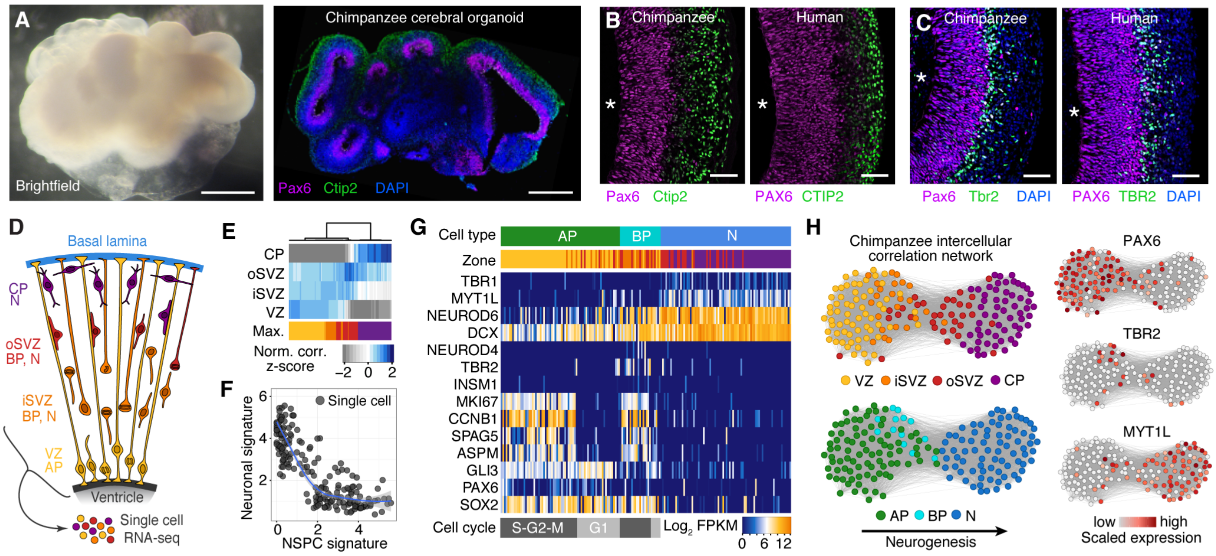


Figure 1: Chimpanzee cerebral organoids recapitulate cortex development. (**A**) Bright-field image showing a representative chimpanzee organoid (Sandra A, left) and a cryosection from a chimpanzee organoid (PR818-5) immunostained for PAX6 (magenta) and Ctip2 (green) combined with DAPI staining (blue) (right) at day 52. Scale bars, 200 µm. (B, C) Cryosections of cortical regions from chimpanzee (Sandra A) and human (SC102A-1) organoids at day 52 immunostained for PAX6 (magenta), Ctip2 (B, green) and TBR2 (C, green), without (**B**) and with (**C**) DAPI staining (blue). Asterisks, ventricular lumen; scale bars, 50 µm. (**D**) Cartoon showing NSPC types (APs, BPs) and neurons enriched in zones within the neocortex at mid-neurogenesis. CP, cortical plate; N, neuron. (**E**) Heatmap showing normalized correlation (Z-score) of single-cell transcriptomes from chimpanzee cerebral organoid cortex with bulk RNA-seq data from laser-microdissected zones [31] from 13 wpc human neocortex. CP, cortical plate. (**F**) Scatterplot showing NSPC and neuronal signature scores derived from analysis of fetal cerebral cortex single-cell transcriptomes (Figure 1—figure supplement 1) calculated for each chimpanzee cerebral organoid cortical cell. (**G**) Heatmap showing expression of AP, BP, and neuron (N) marker genes. Each column represents a single cell, each row a gene. Cell type and maximum correlation to bulk RNA-seq data from cortical zones are shown in the top sidebar. APs and BPs were sub-classified based on G1 (light grey) or S-G2-M (dark grey) phases of the cell cycle. (**H**) Lineage network based on pairwise correlations between chimpanzee cerebral organoid cortical cells reveals a structured topology where VZ-APs connect to cortical plate (CP) neurons (N) through SVZ-BPs. Cells are coloured based on cortical zone (top left) or cell type assignment (bottom left). APs, BPs, and neurons were classified based on maximum correlation with single-cell transcriptomes from the human fetal neocortex. Expression of markers PAX6, TBR2, and MYT1L are shown to the right.

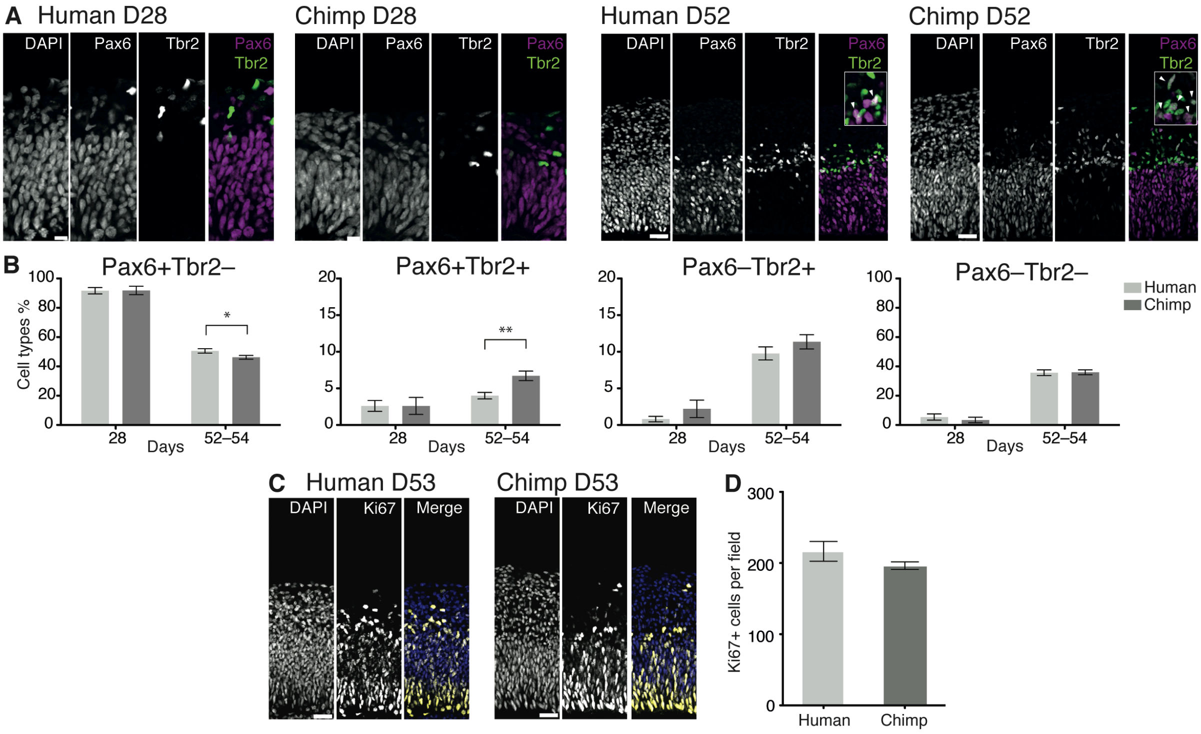


Figure 2: Changes in the proportion of cortical NSPC subtypes and neurons during human and chimpanzee cerebral organoid development. (**A**) Cryosections of cortical regions from human and chimpanzee organoids at day 28 and day 52 immunostained for PAX6 (magenta) and TBR2 (green) combined with DAPI staining. Scale bars; D28, 10 µm; D52, 20 µm. Insets in the D52 merge images show selected areas with PAX6+TBR2+ double–positive nuclei (arrowheads) at higher magnification. (**B**) Quantification of the percentage of PAX6+TBR2–, PAX6+TBR2+, PAX6–TBR2+ and PAX6–TBR2– cortical cells in human (light grey) and chimpanzee (dark grey) organoids at D28 (n = 5 organoids, 50 µm wide field) and D52-D54 (n = 17 organoids, 100 µm wide field). Error bars, SEM; *$p<0.05$, **$p<0.01$. (**C**) Cryosections of cortical regions from human and chimpanzee organoids at D53 immunostained for KI67 (yellow) combined with DAPI staining (blue). Scale bars, 20 µm. (**D**) Quantification of KI67+ cells in a 100 µm wide field in human and chimpanzee organoids at D52-D53 (n = 7). Error bars, SEM

BPs. These lineage relationships were corroborated using a minimal spanning tree algorithm (Figure 1—figure supplement 3G) [251]. Together, these data allowed reconstruction of the chimpanzee organoid cerebral cortex from single-cell transcriptomes.

Chimpanzee and human gene expression in the developing cerebral cortex

To further explore transcriptome similarities and differences between chimpanzee and human cerebral cortex cells, we compared them to the single-cell transcriptomes of 220 fetal human cortex cells (12–13 weeks post-conception (wpc), published in [114], GSE75140) and 207 cortex-like cells from human cerebral organoids (40–80 days, 155 single-cell transcriptomes published in [114], GSE75140; 52 single-cell transcriptomes acquired as part of this study) (Figure 3—source data 1). In a PCA, the first principal component (PC1) separated NSPCs and neurons, whereas PC2 separated species (Figure 3A). Hierarchical clustering of organoid and fetal cells showed that human and chimpanzee organoid and human fetal cells were distributed together within the two main sub-clusters representing NSPCs and neurons (not shown), and showed highly correlated expression of marker gene patterns (Figure 3B).

We constructed an intercellular correlation network, which revealed a VZ sub-network of human and chimpanzees APs that link through BPs expressing iSVZ and oSVZ markers to cortical plate neurons. Generally, APs, BPs, and neurons from human and chimpanzee intermixed, confirming that cells in the chimpanzee organoid cortices have a zonal organization consistent with what is observed histologically (Figure 3C,D). In conclusion, the major proportion of the variation in these data is not between in vitro and in vivo tissues or between species, but among cell states during neurogenesis, confirming that the major features of the genetic programs regulating the NSPC-to-neuron lineage are conserved between human and chimpanzees, and are recapitulated in cerebral organoids.

To identify genes differentially expressed between chimpanzee and human cortex-like cells, we remapped all single-cell transcriptome reads to a consensus human-chimpanzee genome and used human annotations to identify 1-to-1 orthologous genes. We then used a Bayesian approach to identify differentially expressed genes by comparing cerebral organoid APs and neurons between species (ignoring BPs due to the low number of BPs identified). We identified 297 and 279 genes that were more highly expressed in human APs and neurons, respectively, and 283 and 314 genes that were more highly expressed in chimpanzee APs and neurons, respectively (Figure 3E, Figure 3-source data 2). In addition to the between-species comparisons, we identified genes differentially expressed between human or chimpanzee APs and neurons to identify cell-type specific genes (for human: 1328 AP-specific, 1132 neuron-specific; for chimpanzee: 1501 AP-specific, 1166 neuron-specific). The vast majority (94%) of genes that are AP-specific and neuron-specific in humans are not differential between human and chimpanzee (Figure 3E, Figure 3—figure supplement 1). Of the differentially expressed genes

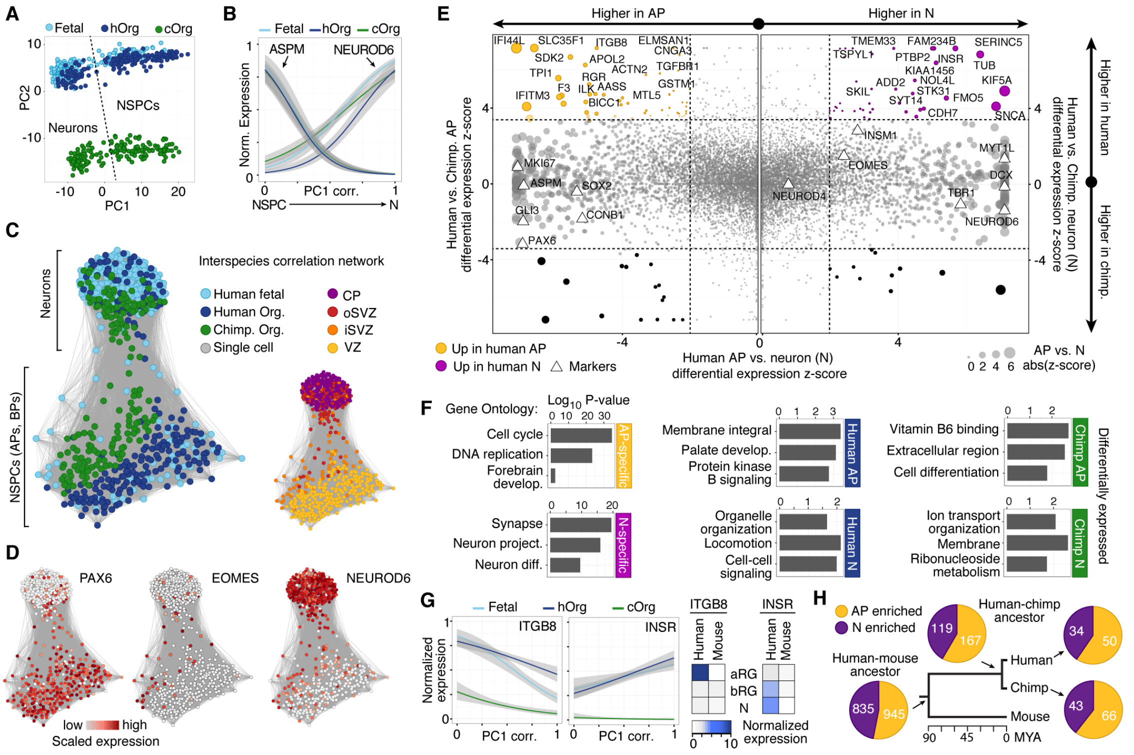


Figure 3: Comparing human and chimpanzee cerebral cortex gene expression. (A) PC1 and PC2 from PCA separated NSPCs and neurons, and human and chimpanzee, respectively. PCA was performed on all single-cell transcriptomes using genes expressed in more than two cells and with a non-zero variance. (**B**) Quasibinomial fit line of representative marker gene expression across cells ordered by correlation with PC1. (**C**) Lineage network based on pairwise correlations between human fetal, human organoid, and chimpanzee organoid cells reveals a differentiation topology from VZ APs through BPs in iSVZ and oSVZ, to cortical plate (CP) neurons, with inter-species mixing in all stages. (**D**) Lineage network (see (**C**)) coloured by scaled expression level of marker genes. (**E**) Scatterplots showing z-scored significance estimates from single-cell differential expression (SCDE) analysis based on Bayesian probabilistic models. Reads from human and chimpanzee were mapped to a consensus genome, and human gene annotations were used for expression counting. The x-axis represents SCDE between human organoid APs vs. human organoid neurons. The y-axes on the left and right plots represent SCDE between human and chimpanzee APs and neurons (N), respectively. Genes coloured as white triangles represent marker genes from Figure 1 and are generally not differentially expressed between human and chimpanzee, but do vary between APs and neurons, validating the SCDE analysis. Yellow and purple circles represent genes upregulated specifically in human APs and neurons, respectively. Circles are sized based on differential expression between human APs and neurons. Figure 3—figure supplement 1 shows a similar plot from the chimpanzee perspective. (**F**) Gene ontology enrichments (-$\log_{10}$ P-value) for differentially expressed gene groups shown in panel E. Left, human APs (yellow) and neurons (N, purple) that are not differential between human and chimpanzee. Center, upregulated in human APs (top) or neurons (N, bottom) compared to chimpanzee. Right, upregulated in chimpanzee APs (top) or neurons (N, bottom) from Figure 3—figure supplement 1. (**G**) Left, expression profiles of ITGB8 and INSR are shown from human organoid, chimpanzee organoid, and human fetal cells ordered by correlation with PC1. Right, bulk RNA-seq data from sorted aRG, bRG, and neurons (N) from human and mouse developing neocortex [62] confirms enriched expression of ITGB8 and INSR in human APs and neurons, respectively. (**H**) The same bulk RNA-seq data was used to confirm and estimate the origin of differential gene expression in APs versus neurons from single-cell organoid data. Pie chart shows the proportion of AP-enriched (yellow) or neuron-enriched (N, purple) genes that are observed in human, chimpanzee, and mouse. Pie charts also show the proportion of genes differential between APs and neurons that are observed only in human and chimpanzee, but not mouse (human-chimp ancestor), or genes specific to human or chimpanzee.

between species, we identified 93 genes that are strongly upregulated in human organoid APs and 72 genes upregulated in human organoid neurons. Gene ontology enrichments suggest that the proteins encoded by some of these genes are integral to cell membranes and involved in intercellular signalling (Figure 3F, Figure 3—source data 2), for example integrin beta 8 (ITGB8) in APs and insulin receptor (INSR) in human neurons. This upregulation of ITGB8 specific to human APs and INSR specific to human neurons is also observed in comparisons between human and mouse [62] (Figure 3G).

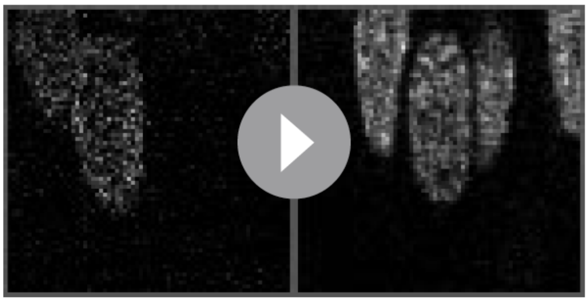
When comparing these results to bulk RNA-seq data from mouse APs and neurons [62], we find that 75% of the genes with expression specific to APs or neurons in humans are also specific to each cell type in the mouse, suggesting that these gene expression programs were already established and likely present in the common ancestor of mouse, human and chimpanzee some 90 million years ago (Figure 3F). Notably, a similar proportion of AP- and neuron-specific genes were gained on the chimpanzee and human branch subsequent to their separation, suggesting that our analysis did not have a strong human bias. About 12% of these genes specific to AP or neurons in both human and chimpanzee were not specific to these cell types in the mouse [62], suggesting that they may be involved in developmental processes specific to the primate cerebral cortex.

Live imaging of NSPC mitoses in human fetal neocortex and cerebral organoids

We used an established live imaging method [247] to compare dividing cortical APs, i.e. cells undergoing mitosis at the ventricular surface (presumably mostly aRG), in slice cultures of both 11–13 wpc human fetal neocortex and human D30 cerebral organoids. We did not observe signs of strong perturbation during live image acquisition in either system, such as mitotic arrest (Figure 4A,C,E; see also Figure 5A–C and Video 1) or lack of nuclear movements and cell death. Chromosome dynamics and spindle orientation of APs, as revealed by the orientation of the metaphase plate, were similar in human developing neocortex and human organoids, both before anaphase (Figure 4A–D,G) and during anaphase (Figure 4A–D,H,I), when cell cleavage initiates. This strongly suggests that cerebral organoids are a suitable model to study live NSPC division and spindle orientation dynamics.

Spindle orientation dynamics are similar in human and chimpanzee NSPCs

Spindle orientation can determine symmetric vs. asymmetric NSPC division [245–248] and is therefore a major candidate mechanism to explain the approximately 3-fold expansion of the neocortex in humans compared to great apes. We compared spindle orientation dynamics between human and chimpanzee APs in cerebral organoids. However, our data revealed no clear differences in spindle orientation, either during metaphase (Figure 4C–G) or shortly after anaphase onset (Figure 4C–F, I–J). As deduced from the orientation of the chromosome



Video 1: Differences in prometaphase-metaphase length between APs of human and chimpanzee cerebral organoids. Related to Figure 5B and C Live tissue imaging of mitotic phases, as reported by chromosomes, in organotypic slice culture of cerebral organoids. Time-lapse is ~1.1 min. Datasets are the same as in Figure 5B and C. Left side: APs in a slice of a D30 human cerebral organoid from iPSC line SC102A-1. Right side: APs in a slice of a D30 chimpanzee cerebral organoid from iPSC line Sandra A. Growing colour bars at the bottom indicate time progression of the respective dividing AP and are synchronized to the beginning of prometaphase (in green). Metaphase plate time is in yellow and anaphase time is in red. Note the slower progression of the dividing human AP on the left. DOI: 10.7554/eLife.18683.021.

plates, most APs in both human and chimpanzee divided with a cleavage orientation largely perpendicular to the apical, ventricular surface, showing deviations of fewer than 30° from a perfect orthogonal cleavage. Oblique and near-horizontal orientations were also observed, but at a much lower abundance and at similar frequencies in chimpanzee and human organoids (Figure 4H–J). This shows that the frequency of asymmetric cell division caused by oblique spindle orientation is most likely not a major difference between human and chimpanzee APs.

Longer prometaphase-metaphase in human than great ape APs

We noticed, however, unexpected differences between human and chimpanzee APs in their progression through mitosis. Specifically, measurement of the length of the various phases of mitosis (for details, see Materials and methods) revealed that APs in 11–13 wpc fetal human neocortex and D30 cerebral organoids remained approximately 5 min longer in prometaphase-metaphase than APs in chimpanzee organoids (Figure 5A–C,E; Video 1). By comparison, prometaphase-metaphase of APs in slice culture of mouse neocortex, a well-characterized model system for neurogenesis, lasted for only approximately half the amount of time than human APs (Figure 5D,E; Figure 5—source data 1).

To trace the specific phase of mitosis when this difference arises, we used chromosome morphology and dynamics to determine the time chromosomes spent congressing toward the equatorial plane of the cell (defined here as 'prometaphase') and the time they spent tightly aligned as a metaphase plate (defined here as 'metaphase'). Remarkably, the longer prometaphase-metaphase of human than chimpanzee APs was specifically due to a 40-60% lengthening of metaphase (Figure 5A–C,G), whereas prometaphase was not significantly different (Figure 5A–C,F; Video 1). By contrast, in mouse APs, both prometaphase and metaphase were found

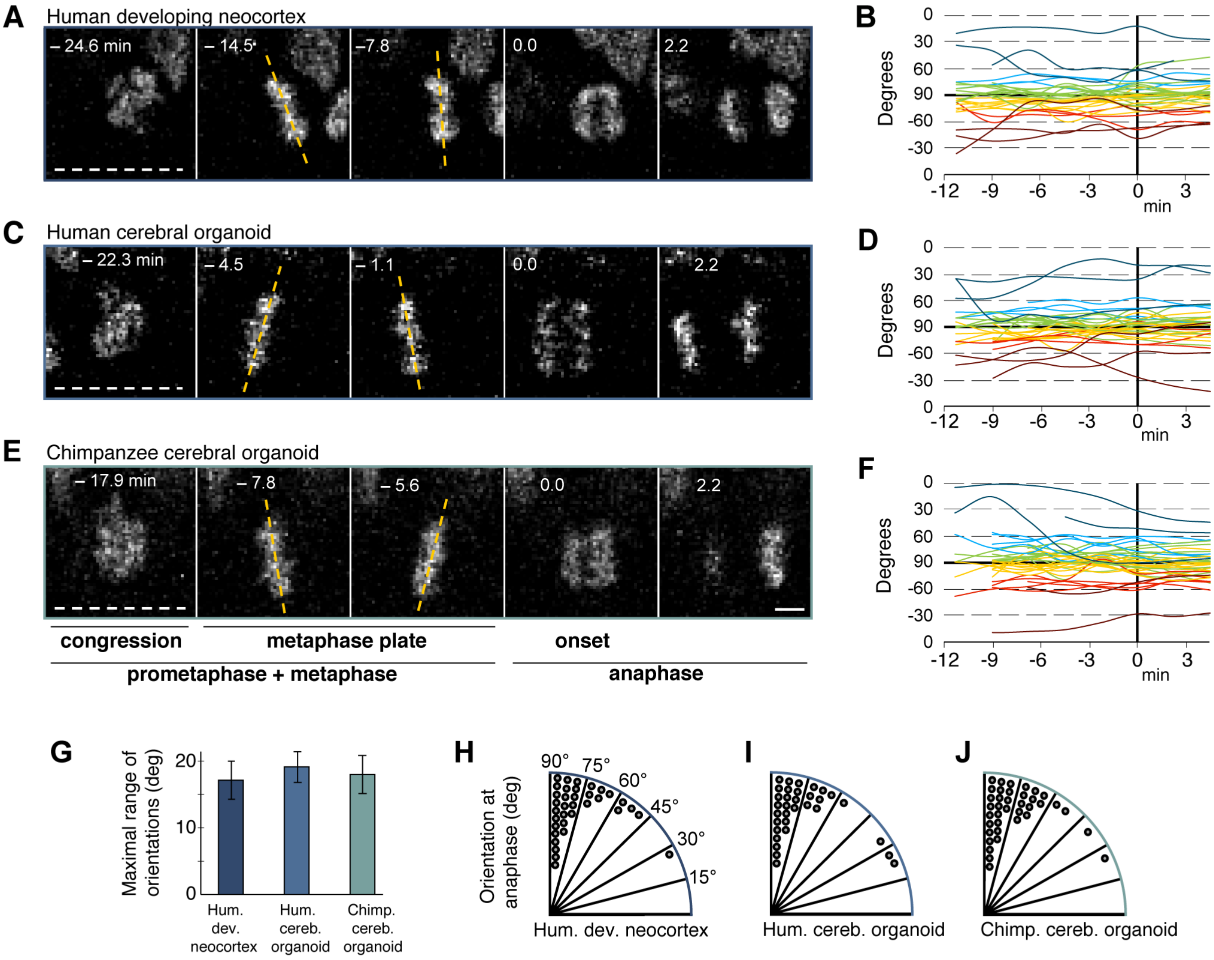


Figure 4: Spindle orientation variability is similar between APs of human developing neocortex, human organoids and chimpanzee organoids. Live tissue imaging of spindle orientation, as reported by chromosome plate orientation, in organotypic slice culture of developing neocortex and cerebral organoids. Measurements were started after all chromosomes had formed a tight metaphase plate. 0 min is anaphase onset. Time-lapse is ~ 1.1 min. (**A, C, E**) APs in a coronal slice of 13 wpc human frontal neocortex (**A**), in a slice of a D30 human cerebral organoid from iPSC line SC102A1 (**C**), and in a slice of a D30 chimpanzee cerebral organoid from iPSC line Sandra A (**E**). The time indicated on each image is when that image was taken, relative to anaphase onset (0 min). White dashed lines, ventricular surface. Yellow dashed lines indicate the two metaphase plate orientations with the greatest difference to each other. Scale bar, 5 µm. (**B, D, F**) Quantification of all orientations of the chromosome plates from the beginning of the metaphase plate stage to anaphase, for APs in the three respective tissues described in (**A, C, E**). To facilitate tracing, individual tracks are colour-coded acaccording to the initial range of the track, and the 90°-0° range is depicted twice (green and yellow, 90°-75°; cyan and red, 75°-60°; blue and dark red, 60°-0°; 90° indicates perfectly vertical chromosome plates). (**G**) Maximal range of chromosome plate orientations for APs, from the beginning of the metaphase plate stage to anaphase onset, as determined in the measurements shown in (**B, D, F**). Data are the mean ± SEM of 34 ≥ APs from 3 independent experiments each. (**H, I, J**) Orientation of chromosome plates at 2.2 min after anaphase onset, which indicates the predicted plane of cleavage, as determined in the measurements shown in (**B, D, F**). 90° indicates a perfectly vertical cleavage plane.

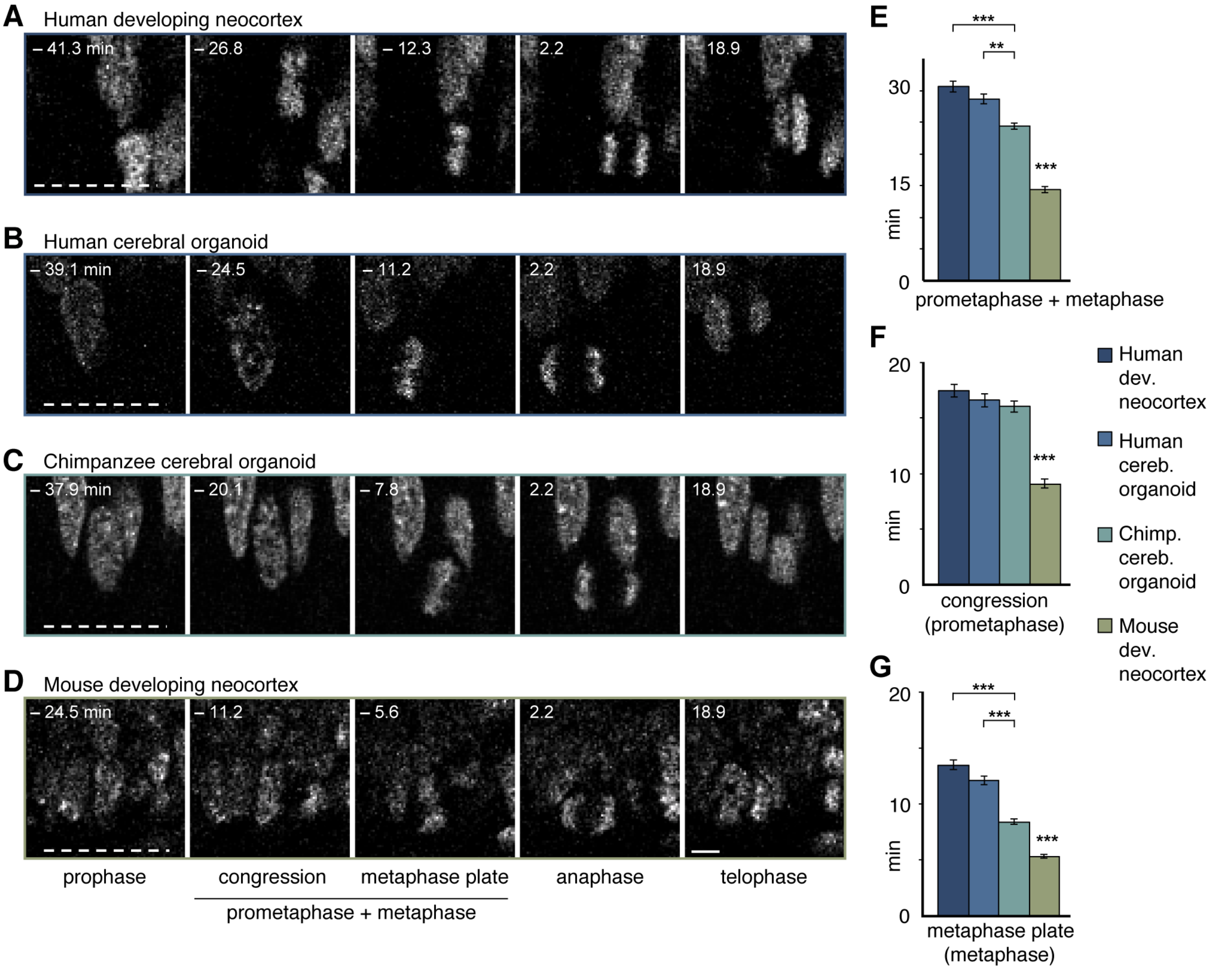


Figure 5: Differences in prometaphase-metaphase length between APs of human developing neocortex, human organoids, chimpanzee organoids and mouse developing neocortex. Live tissue imaging of mitotic phases, as reported by chromosomes, in organotypic slice culture of developing neocortex and cerebral organoids. 0 min is anaphase onset. Time-lapse is ∼1.1 min. (**A**–**D**) APs in a coronal slice of 13 wpc human frontal neocortex (**A**), in a slice of a D30 human cerebral organoid from iPSC line SC102A-1 (**B**), in a slice of a D30 chimpanzee cerebral organoid from iPSC line Sandra A (**C**), and in a coronal slice of E14.5 mouse neocortex. The time indicated on each image is when that image was taken, relative to anaphase onset (0 min). White dashed lines, ventricular surface. Scale bar, 5 µm. (**E**–**G**) Time between the start of chromosome congression and anaphase onset (referred to as 'prometaphase + metaphase') (**E**), between the start of chromosome congression and the formation of a metaphase plate (referred to as 'prometaphase') (**F**), and between the formation of a metaphase plate and anaphase onset (referred to as 'metaphase') (**G**), for APs in the four tissues described in (**A**–**D**). Data include APs from 11–13 wpc human neocortex, organoids from the human iPSC lines SC102A-1 and 409b2, and chimpanzee iPSC lines Sandra A and PR818-5, and are the mean ± SEM of $\geq$60 APs from 4 independent experiments each. Bracket with **<0.01; brackets with ***p<0.001; ***p<0.001 (mouse vs. all primate tissues).

to be significantly shorter than the respective mitotic phases in human and chimpanzee APs (Figure 5D,F,G; Figure 5—source data 1).

None of the other mitotic phases (prophase, anaphase, telophase) differed in length betweenAPs in human fetal neocortex and human cerebral organoids vs. chimpanzee organoids. However, anaphase of mouse APs was found to be significantly shorter than that of human and chimpanzee APs (Figure 5—figure supplement 1A; Figure 5—source data 1). These differences between species in the individual mitotic phases were reflected in the cumulative length of total mitosis, which was significantly shorter in mouse APs than human and chimpanzee APs (Figure 5—figure supplement 1B).

To search for potential functional implications of these observations, we next quantified and compared the length of prometaphase-metaphase in human and chimpanzee APs of day 52 (D52) cerebral organoids, and compared the results with those of D30 organoids. Prometaphase-metaphase (Figure 5—figure supplement 2A) and metaphase alone (Figure 5—figure supplement 2C; Figure 5—source data 1) were shorter in D52 than in D30 human APs, and not anymore statistically significantly different in length from D52 chimpanzee APs. The longer metaphase of human than chimpanzee organoid APs may therefore characterise early phases of cortical development, when proliferative AP divisions are predominant.

We also generated cerebral organoids from an orangutan iPSC line and determined the length of AP prometaphase-metaphase. This revealed that the length of prometaphase-metaphase in orangutan D30 organoid APs was similar to that of chimpanzee APs and significantly shorter than that of human organoid APs (Figure 5—figure supplement 3A,B). As was the case for the human-chimpanzee AP comparison, the shorter prometaphase-metaphase of orangutan than human APs was due to a shorter metaphase (Figure 5—figure supplement 3A,D) rather than prometaphase (Figure 5—figure supplement 3A,C; Figure 5—source data 1). Together, these data indicate that human APs specifically lengthen prometaphase-metaphase as compared to great ape APs.

In light of these differences in the duration of mitotic phases, it was of interest to compare the length of the total cell cycle of human and chimpanzee organoid APs. Using cumulative EdU labelling of D52-D54 cerebral organoids (Figure 5—figure supplement 4A), we found a relatively minor (6%) difference in total cell cycle length, with human APs (PAX6+TBR2– cells) exhibiting a 2.7 hr longer cell cycle (46.5 h) than chimpanzee APs (43.8 h) (Figure 5—figure supplement 4B). However, a notable difference between the two species was the length of S-phase, which was nearly 5 hr longer in human (17.5 h) than chimpanzee (12.8 h) organoid APs (Figure 5—figure supplement 4B).

The prometaphase-metaphase lengthening in humans occurs upon neural differentiation

To investigate the origin of the longer metaphase in human than chimpanzee APs, we measured mitotic phase lengths in the original iPSCs used to grow the cerebral organoids. This revealed that both the human and chimpanzee organoid APs had a longer prometaphase-metaphase than their respective iPSCs of origin, showing that this general lengthening was due to the transition between iPSCs and the organoids of both species (Figure 6A,B,E). In human APs, however, the lengthening was greater than in chimpanzee APs. In contrast to APs, human and chimpanzee iPSCs had similar prometaphase-metaphase lengths (Figure 6A,B,E; Figure 5—source data 1).

Further dissection into individual phases revealed that, whereas both human and chimpanzee APs had a longer prometaphase than their iPSCs of origin (Figure 6A,B,F), only human APs had a longer metaphase when compared to the iPSCs of origin (Figure 6A,B,G; Figure 5—source data 1).This shows that prometaphase-metaphase lengthened in both species as APs were generated during cerebral organoid formation with the accompanying neural differentiation. However, the lengthening characteristics were species-specific. The lengthening was greater in humans than chimpanzees because the metaphase plate stage became longer only in human APs.

To determine whether prometaphase-metaphase length may differ between chimpanzees and humans also in another cell type, we measured mitotic phases in human and chimpanzee B cells. In contrast to fetal tissue, these cells can be obtained not only from humans but also chimpanzees by collecting blood, that is, without major invasive procedures. The length of prometaphase-metaphase in B cells, as well as prometaphase and metaphase measured individually, were similar to that in iPSCs (Figure 6 C–G), and hence significantly shorter than in human or chimpanzee APs. By contrast, the other mitotic phases were similar among organoid APs, iPSCs and B cells in both species (Figure 6—figure supplement 1; Figure 5—source data 1). This raises the intriguing possibility that lengthening of prometaphase-metaphase could be specific to ape and human NSPCs and, furthermore, that lengthening of the metaphase plate time could be specific to human NSPCs.

Longer prometaphase-metaphase in proliferative than neurogenic mouse APs

To investigate potential functions of prometaphase-metaphase lengthening, we asked whether mitotic phases were different between proliferating and neurogenic APs. To this end, we measured mitotic phase lengths in a transgenic mouse line where EGFP is expressed under the promoter of the pan-neurogenic marker Tis21 in neurogenic but not proliferative NSPCs [8, 252]. This revealed that prometaphase-metaphase was longer in proliferative AP divisions

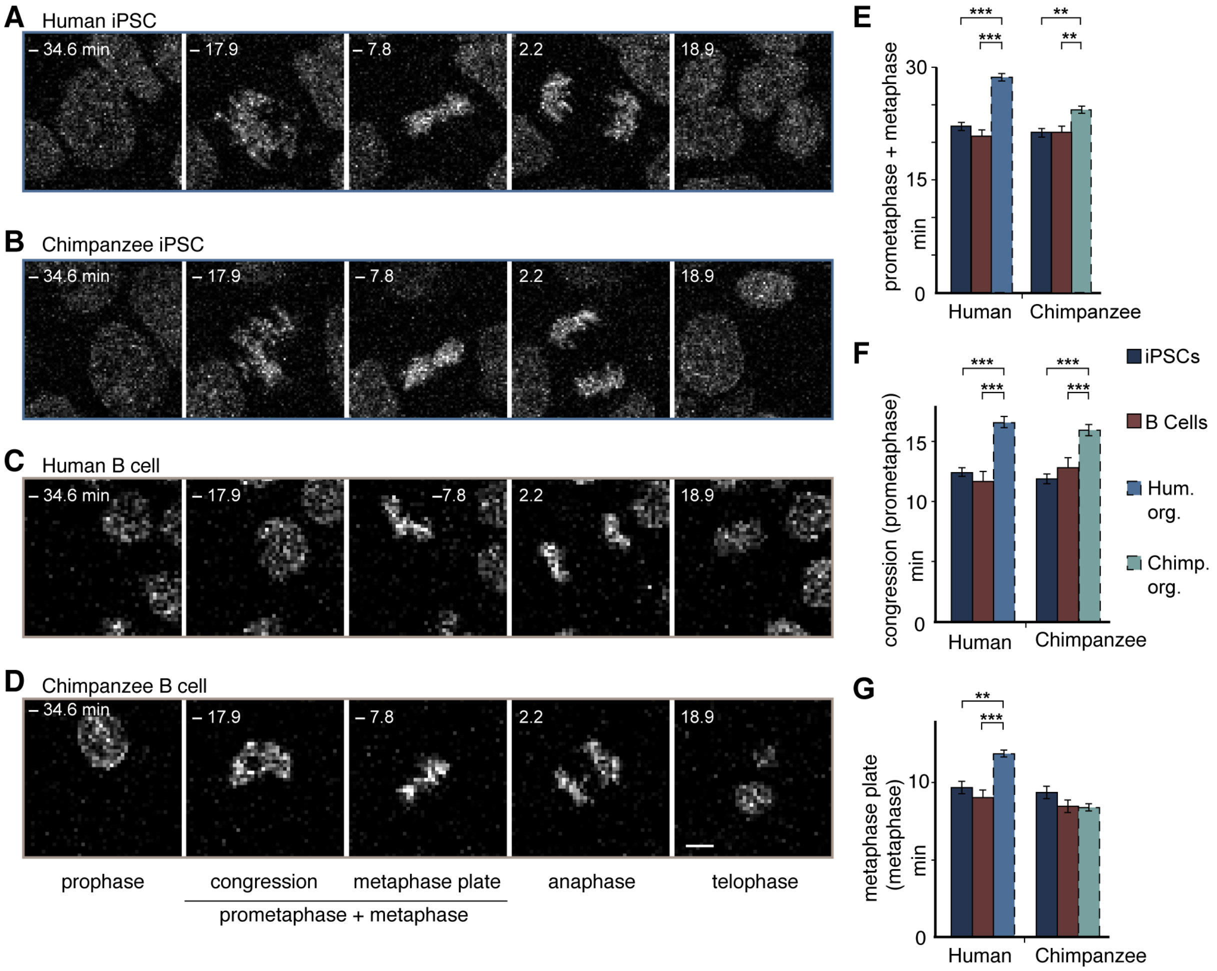


Figure 6: Human and chimpanzee organoid APs exhibit longer prometaphase, and human organoid APs longer metaphase, than their iPSC lines of origin or B cells. Live imaging of mitotic phases, as reported by chromosomes, in human and chimpanzee iPSCs and B cells. 0 min is anaphase onset. Time-lapse is $\sim$1.1 min. (**A–D**) Human iPSC (SC102A-1) (**A**), chimpanzee iPSC (Sandra A) (**B**), human B cell (A158) (**C**), and chimpanzee B cell (Dorien) (**D**). The time indicated on each image is when that image was taken, relative to anaphase onset (0 min). Scale bar, 5 µm. (E–G) Time between the start of chromosome congression and anaphase onset (referred to as 'prometaphase + metaphase') (**E**), between the start of chromosome congression and the formation of a metaphase plate (referred to as 'prometaphase') (**F**), and between the formation of a metaphase plate and anaphase onset (referred to as 'metaphase') (**G**). Data include cells from each of the following iPSC lines: human, SC102A-1 and 409b2; chimpanzee, Sandra A and PR818-5; and from the following B cell lines: human, A144, A156 and A158; chimpanzee, Jahaga, Ulla and Dorien. For comparison, the relevant mitotic phase lengths of human and chimpanzee cerebral organoid APs from Figure 5 are shown (columns with dashed line). Data are the mean $\pm$ SEM of $\geq$30 cells from $\geq$ independent experiments each. **p<0.01; ***p<0.001.

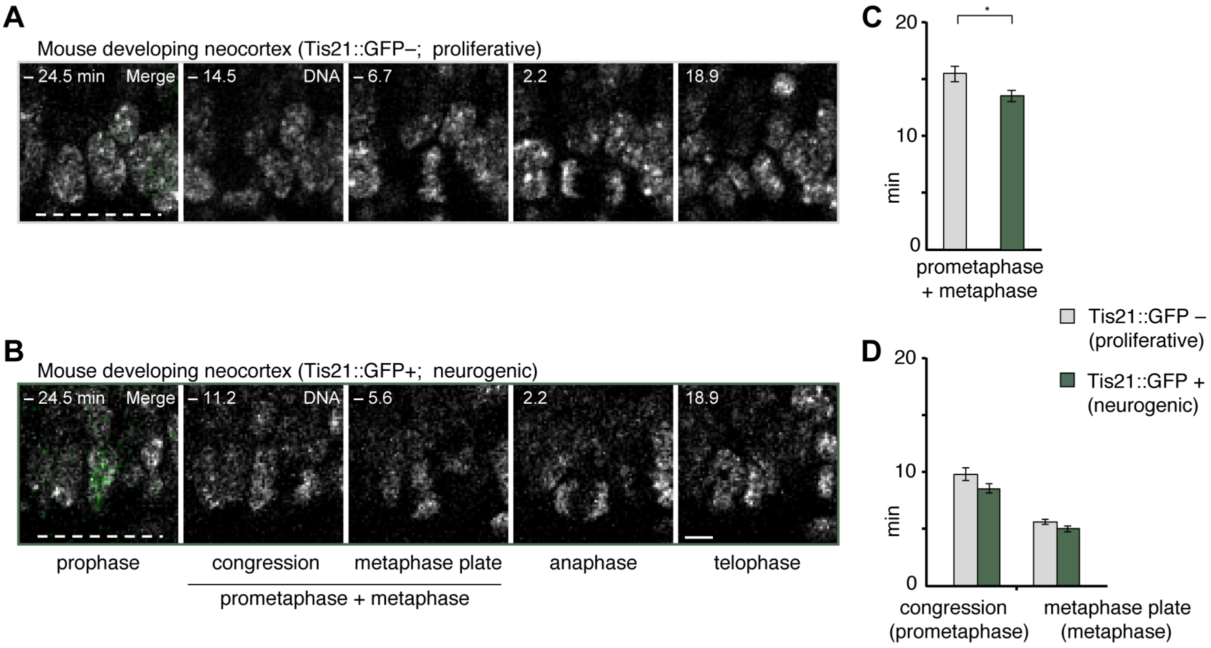


Figure 7: Prometaphase-metaphase is longer in proliferative than neurogenic mouse APs. Live tissue imaging of mitotic phases, as reported by chromosomes, in organotypic slice culture of E14.5 mouse neocortex. 0 min is anaphase onset. Time-lapse is ∼1.1 min. Data is from the same mouse dataset shown in Figure 5, but distinguishes between Tis21::GFP– (proliferative) and Tis21::GFP (neurogenic) APs. (**A,B**) APs in a coronal slice of mouse E14.5 dorsolateral telencephalon, either negative (**A**) or positive (**B**) for expression of Tis21::GFP. The time indicated on each image is when that image was taken, relative to anaphase onset (0 min). White dashed lines, ventricular surface. Scale bar, 5 µm. Image panels in (**B**) are the same as in Figure 5D, but the Tis21::GFP fluorescence (green) is included in the prophase image (merge). The GFP channel is also merged in the prophase image of (A), and the other panels are DNA staining only. (C,D) Length of prometaphase and/or metaphase in proliferative vs. neurogenic APs. Data are the mean ± SEM of 41 Tis21::GFP– and 37 Tis21::GFP APs from 4 independent experiments. *$p<0.05$. (**C**) Time between the start of chromosome congression and anaphase onset (referred to as 'prometaphase + metaphase'). (**D**) Time between the start of chromosome congression and the formation of a metaphase plate (referred to as 'prometaphase', left), and time between the formation of a metaphase plate and anaphase onset (referred to as 'metaphase', right).

(Tis21–) than in neurogenic AP divisions (Tis21+), whereas the separate phases were not significantly different (Figure 7; Figure 5—source data 1). These results suggest that a lengthening of prometaphase-metaphase may be characteristic of proliferating NSPCs.

Gene expression in human and chimpanzee mitotic APs

We used the single-cell RNA-seq data to identify organoid APs in different phases of the cell cycle (Figure 8A, Figure 8—figure supplement 1) and searched for genes that might be involved in human-specific lengthening of the metaphase. We compared human organoid APs in G1 with APs in G2-M and identified 395 genes with enriched expression in G2-M (Figure 8B). We next compared human APs in G2-M with human iPSCs (TkDA3-4) and an endothelial cell line (HUVEC; both single-cell RNA-seq data sets in GSE81252) to understand the specificity of

G2-M regulation in APs. We found that nearly all genes upregulated in human APs in G2-M compared with human APs in G1 were also upregulated during G2-M in iPSCs and endothelial cells (Figure 8C). Therefore, the expression level of these genes is unlikely to contribute to the specificity of mitotic control of human APs in G2-M. However, we identified many genes that were highly expressed throughout the human AP cell cycle and were specific to APs. Genes with the highest specificity score encoded canonical cerebral cortex patterning transcription factors such as PAX6, ID4, and GLI3, as well as proteins involved in cell adhesion and ECM signalling (CDH4, EFNB1/2, COL4A2). Notably, no genes associated with cell cycle, kinetochore, or spindle terms were specific to human APs (Figure 8C, inset). Of genes specific to APs, a subset were also differentially expressed between human and chimpanzee cerebral organoids (APOLD1, BICC1, EFNB1, GSTM1, IFI44L, ITGB8, SDK2, SEMA5A, SLC35F1, ZNF516), which makes them candidates for the unique regulation of AP proliferation in humans (Figure 8D).

Discussion

We have characterized cerebral organoids generated from chimpanzee iPSCs, including a newly generated iPSC line, and shown that their cytoarchitecture, cell type composition, and neurogenic gene expression programs are remarkably similar to human cerebral organoids and to human fetal neocortex. This extends a very recent study [59] and establishes cerebral organoids as a valid system to compare human and chimpanzee NSPC behaviour. Using this system, we have shown that human and chimpanzee APs differ in that prometaphase-metaphase is longer in humans than in chimpanzees. This difference was also observed between human and orangutan and reflects a greater extent of prometaphase-metaphase lengthening that occurs as human APs are generated during cerebral organoid development from IPSCs. There are two intriguing implications as to the biological significance of this prometaphase-metaphase lengthening in human APs.

One is related to the fate of the progeny arising from AP division. Mouse Tis21::GFP-negative APs, which are known to undergo proliferative divisions to generate more APs, have a longer prometaphase-metaphase than Tis21::GFP-positive APs, which are known to undergo neurogenic divisions to generate BPs [8]. The longer prometaphase-metaphase in human than chimpanzee APs would therefore be consistent with a greater tendency for proliferative than neurogenic divisions. In this respect, other changes in progeny fate have also been recently observed in a different context, upon an experimentally induced and considerable prolongation of AP mitosis in embryonic mouse neocortex [253].

Another set of observations are consistent with the notion that the longer prometaphase-metaphase in human than chimpanzee APs may indicate a greater tendency for proliferative

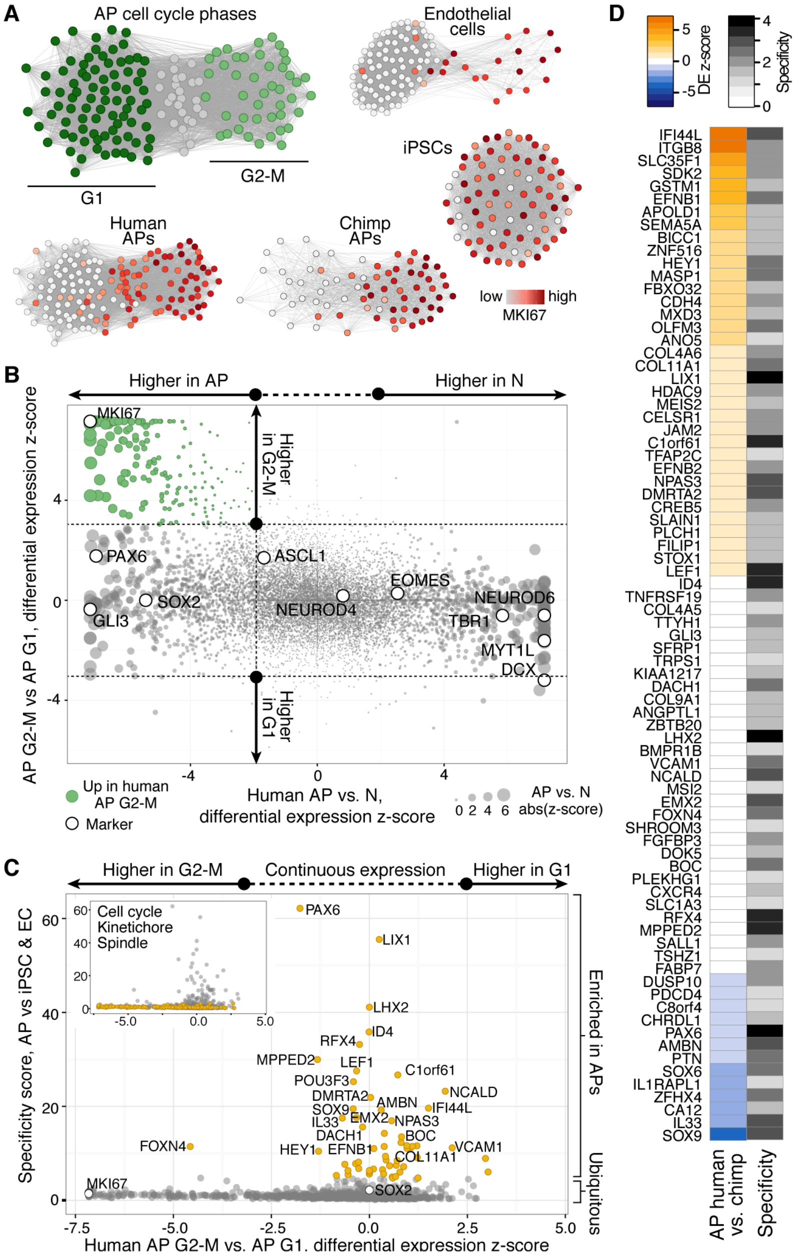


Figure 8: Differential gene expression during AP mitotic phases. (**A**) G1 (dark green) and G2-M (light green) cell cycle phases were assigned to cells by performing PCA using genes involved in cell cycle regulation. PC1 and PC2 described cell cycle phases, and the top 50 correlating and anticorrelating genes were used to infer an intercellular correlation network for human and chimp APs, human iPSCs, and a human endothelial cell line. Networks are coloured based on the expression level of MKI67. (**B**) Scatterplot shows z-scored significance estimates from single-cell differential expression (SCDE) analysis between human organoid APs vs. neurons (N, x-axis) and APs in G2-M vs. APs in G1 (y-axis). Genes coloured as white circles represent marker genes and green circles represent genes upregulated specifically in APs in G2-M. Circles are sized based on differential expression between human APs and neurons. (**C**) iPSC and endothelial cell (EC) expression was used to assign a specificity score for genes enriched in human organoid APs compared to neurons (higher in AP genes from panel B). The specificity score is plotted against the differential expression between APs in G2-M and APs in G1. Cells with high AP specificity scores are in yellow in the main scatter plot. This shows that nearly all genes enriched in G2-M phase of the AP cell cycle are not specific to APs, but also enriched in G2-M of mitotic iPSCs and endothelial cells. (**D**) Heatmap shows the differential expression score between human and chimpanzee APs (z-score) and AP specificity score (Log_2 normalized) of the same genes that are specific to APs relative to endothelial cells and iPSCs.

than differentiative divisions. The human vs. chimpanzee prometaphase-metaphase difference decreased in the course of organoid cortical development from D30 to D52, when one would expect proliferative AP divisions to decrease and differentiative AP divisions to increase.

Further support for this notion was obtained by analysis of the interphase of the cell cycle, specifically S-phase. Mouse Tis21::GFP-negative (proliferative) APs have previously been shown to have a longer S-phase than Tis21::GFP-positive (differentiative) APs [254]. The substantially longer S-phase of human than chimpanzee APs observed here is therefore also in line with human APs having a greater tendency for proliferative divisions.

Finally, the changes in the abundance of NSPC types in the course of cerebral organoid develop- ment yielded data supporting a greater AP proliferation in human than chimpanzee. Specifically, the proportion of PAX6+TBR2– NSPCs, located in the VZ and thus constituting proliferating APs, decreased in both human and chimpanzee cerebral organoids, but the value reached in human organoids was slightly higher than that in chimpanzee organoids (Figure 2B). Conversely, the proportion of PAX6+TBR2+ NSPCs, located in the basal VZ and SVZ and constituting BPs with neurogenic potential, showed a greater increase in chimpanzee than human cerebral organoids. In sum, two independent lines of evidence, the detailed analysis of AP mitosis phase lengths and the determination of the proportions of the various NSPC types, support the concept that a longer neurogenic period [255], which in turn implies a longer phase of NSPC proliferation [59], contributes to the greater expansion of the neocortex in humans than the great apes.

The second implication as to the biological significance of the longer prometaphase-metaphase in human than chimpanzee APs concerns the fact that these are the phases of mitosis when chromosomes prepare for segregation, to ensure that only one copy of each chromosome is distributed to each nascent daughter cell [256]. The longer duration of prometaphase-metaphase in human than chimpanzee APs, in particular of the metaphase plate stage (Figure 5B), may therefore reflect some difference between the two species with regard to the preparation for chromosome segregation.

If the longer prometaphase-metaphase in human than chimpanzee APs reflects a greater tendency for proliferative than neurogenic divisions in the human NSPCs, why did we not detect significant differences between human and chimpanzee APs in spindle orientation, a parameter previously shown to affect the mode of AP division [245–248] This may be due to spindle orientation variability between individual APs being greater than potential inter-species differences. This suggests that, in the cell types and stages analysed, spindle orientation may not play a key role in human vs. chimpanzee neurogenesis. Alternatively, this may reflect the fact that differences in proliferative versus neurogenic AP divisions can occur without a change in spindle orientation [246, 257, 258]. In this context, differences between human and chimpanzee NSPCs of relevance for neocortex expansion are likely to be small. Consistent with this view, our single-cell transcriptome analyses revealed only few differences between human and chimpanzee, and the differences in the proportions of organoid NSPC populations were in the range of a few percentage points. Furthermore, the 5 min longer prometaphase-metaphase in human than chimpanzee APs constituted only a fraction of the total duration of their mitosis. These small differences nevertheless provide a set of clues as to which NSPC features may underlie the differential extent of neocortex expansion in humans versus apes, and are consistent with a scenario in which the accumulation of such small differences during evolution may have resulted in the distinct chimpanzee and human neocortices.

Materials and methods

Neocortex tissue

Human fetal brain tissue (11–13 weeks post conception (wpc)) was obtained with informed written maternal consent followed by elective pregnancy termination, and neocortex was dissected at room temperature, as described previously [62]. Research involving human fetal brain was approved by the Ethical Review Committee of the Universitätsklinikum Carl Gustav Carus of the Technische Universität Dresden (reference number EK100052004). In addition, research was approved by the Institutional Review Board of the Max Planck Institute of Molecular Cell Biology and Genetics.

Mice were kept pathogen-free at the Biomedical Services Facility of the MPI-CBG. All mouse embryos were heterozygotes of the Tis21::GFP knock-in line [8]. Imaging was performed in the dorsolateral telencephalon of E14.5 embryos, at a medial position along the rostro-caudal axis. Embryonic day (E) 0.5 was defined as noon of the day of vaginal plug identification. All experiments using mice were performed according to the German Animal Welfare Legislation.

Cell lines and organoid culturing

Two human iPSC lines (409b2, SC102A-1), two chimpanzee iPSC lines (PR818-5, Sandra A), and one orangutan iPSC line (Toba) were used to generate cerebral organoids in this study. 409b2 was purchased from the RIKEN BRC cell bank and SC102A-1 was purchased from System Biosciences. PR818-5 (0818) was obtained as a kind gift from F. Gage [82] from the Salk Institute for Biological Studies (La Jolla, CA). Sandra A and Toba were generated in collaboration with Shinya Yamanaka following a nonviral transfection method [81]. Briefly, blood was collected from a chimpanzee and an orangutan, both housed at the Leipzig Zoo, and leukocytes were isolated by Ficoll gradient centrifugation, which were then used for reprogramming to iPSCs. DNA sequencing revealed no chromosome aberrations, and RNA-seq and immunohistochemistry confirmed pluripotent gene and protein expression signatures. Primate blood samples used to generate iPSCs were obtained by certified veterinarians during annual medical examinations or other necessary medical interventions, meaning that no invasive procedures were performed on primates for the sole purpose of our research project. The Max Planck Institute for Evolutionary Anthropology has an institutional permit for the transport of biological material derived from endangered species (DE216-08, see http://cites.org/common/reg/si/e-si-beg.shtml). Human iPSC line TkDA3-4 [95] was used to generate iPSC single-cell transcriptomes. iPSC lines were cultured under standard iPSC culturing methods on matrigel (BD Biosciences) using mTeSR1 (Stemcell Technologies). Human endothelial cells. (HUVECs, Lonza, Basel, Switzerland) were maintained in endothelial growth medium (EGM) (Lonza) at 37°C in a humidified 5% CO_2 incubator. Single cell transcriptome analysis confirmed the identity of human and chimpanzee iPSCs and human endothelial cells, and showed no contamination with other cell lines. B-cell lines were generated from blood obtained from three human (A144, A156, A158) and three chimpanzee (Dorien, Jahaga, Ulla) individuals. Withdrawal and processing of blood samples was performed according to approved protocols, and was performed for chimpanzee during necessary veterinary interventions. Lymphocytes were isolated from blood using a Ficoll gradient centrifugation. Immortalization was performed by adding Epstein Barr virus (EBV) supernatant to the lymphocytes and further cultivation of the cells until colonies of immortalized B-lymphocytes were established [259]. B-cells were maintained in RPMI with 10% FBS, 1% Glutamax and 2% penicillin/streptomycin. Cell lines were regularly tested for mycoplasma using a PCR-based test (Minerva Biolabs) and found to be negative.

Human and chimpanzee cerebral organoids were generated from the above iPSCs and cultured for the indicated times as described previously for human cerebral organoids [92, 250], with minor modifications [114].

Single-cell RNA-seq experiments

Preparation of single-cell suspensions from cerebral organoids To generate single-cell suspensions, cerebral organoids were either dissociated as a whole or first sliced using a vibratome to dissect cortical regions. Whole organoids were washed three times in PBS and incubated at 37°C in 2 ml Accutase (Sigma) plus 2 U/ml DNAse I (New England Biolabs) for ~45 min. For dissections, organoids were washed using PBS and embedded into 4% low-melting agarose (Sigma) and sliced into 150 µm sections using a vibrating microtome (Ci 7000 smz, Camden Instruments). Slices were kept in differentiation plus vitamin A (Diff +VA) medium [92] and inspected under a stereomicroscope (Leica) to dissect cortical regions. Selected regions were washed three times with PBS and incubated in ~200 µL Accutase with DNAse I at 37°C for ~45 min. Additional mechanical dissociation was performed by triturating the tissue. Subsequently, cells were filtered through a 30 µm cell filter (Miltenyi Biotec), washed with Diff +VA medium and spun down at 300 g for 5 min. The resulting pellet was resuspended in 30-50 µL (for cortical slices) or 250500 µL (for whole organoids) of Diff +VA medium. In case of excessive debris being present, cells were cleaned using a Percoll (Sigma) gradient centrifugation and the resulting pellet was resuspended in 30-50 µL Diff +VA medium. Counting of cells was performed using a Countess automated cell counter (Invitrogen) and by staining with Trypan blue. For single-cell experiments, cell suspensions were diluted to a final concentration of 450–600 cells/µL.

Single-cell cDNA generation These steps were performed as described [114, 161]. Depending on the size distribution of the cells, cells were loaded at a concentration of 250–500 cells per ml onto small (5–10 µm) or medium (10–17 µm) integrated fluidic circuits (IFCs, Fluidigm). Lysis, reverse transcription and amplification were performed on the Fluidigm C1 platform using the SMARTer Ultra Low RNA Kit for the Fluidigm C1 system. External RNA Control Consortium (ERCC) spike-ins (Ambion) were added to the lysis mix at a dilution of 1:80,000. Resulting cDNA was quantified and checked for its size distribution using a capillary gel electrophoresis system (Fragment Analyzer, Advanced Analytical, 1–6000 bps High Sensitivity).

RNA-Seq library preparation and sequencing Each cell's cDNA was diluted and libraries were prepared using Nextera XT DNA library preparation kits (Illumina). Up to 96 single-cell libraries were pooled and cleaned up using solid phase reversible immobilization (SPRI) beads (Thermo Scientific). Quantification and library size distribution was assessed on a Bioanalyzer (Agilent) platform using High Sensitivity DNA chips. Up to 192 cells were pooled and sequenced in 100-bp paired-end mode on one lane of an Illumina HiSeq 2500 platform (rapid mode).

Read processing, mapping and gene quantification Base-calling, adaptor trimming and demultiplexing of reads was performed using a custom pipeline based on freelbis [260], leeHom [261] and deML [262]. Demultiplexed reads were mapped using TopHat v2.0.14, and FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values per gene were quantified using Cufflinks v.2.2.1 [263]. Human reads were mapped to the hg38 reference genome (release 77) and chimpanzee reads were mapped to panTro4 (release 80). The raw FPKM data of all single cells were combined into one master table and transformed to log2 (FPKM +1). (RStudio Team, 2015) was used to run [264], scripts to perform principal component analysis (PCA, FactoMineR package), hierarchical clustering (stats package), differential expression analysis (SCDE package), and to construct heatmaps, scatter and line plots, dendrograms, bar graphs, pie charts and histograms. Generally, ggplot2 and gplots packages were used to visualize the data. Gene ontology enrichment analyses were performed using DAVID informatics Resources 6.7 of the National Institute of Allergy and Infectious Diseases [265].

Analysis of chimpanzee single-cell RNA-seq data The Seurat package [166] implemented in R was used to identify cell populations present in chimpanzee organoids (Figure 1—figure supplement 2). T-distributed stochastic neighbour embedding (tSNE) was performed on all chimpanzee organoid cells using the most significant genes (p-value $<$10–3, with a maximum of 200 genes per principal component) that define the first 6 principal components of a PCA analysis on the data set. In Figure 1E we calculated for each chimpanzee organoid cortex cell the Spearman correlation of its transcriptome (all genes) with bulk transcriptome data from each of 4 microdissected human cortical zones (VZ, iSVZ, oSVZ, CP, mean expression value of each gene across 4 replicates from 13 weeks post conception, data published in ([31] GSE38805). We hierarchically clustered (Pearson's correlation distance metric) cells based on their correlation coefficient with germinal zones and visualized the clustering in a heatmap showing correlation coefficients scaled across zones (mean-centering and dividing by standard deviation). The scaling enables a better comparison between cells, since the maximum and minimum correlation for each cell is color-coded in the same way after scaling. We used this analysis to identify the zone with which each individual cell had a maximum correlation.

NSPC and neuron signatures (Figure 1—figure supplement 3, Figure 1F) were defined by the top 100 genes correlating or anti-correlating with PC1 from PCA of human fetal neocortex, respectively. Each fetal, human organoid, and chimpanzee organoid cortex cell was scored for the NSPC or neuron signature by summing the number of genes from each signature that have an expression greater than log2 FPKM of 5, and normalizing by the number of all genes expressed above log2 FPKM of 5 for each cell. Lineage network analysis and visualizations were done using igraph implemented in R (http://igraph.sf.net). To construct the chimpanzee cellular network, we computed a pairwise correlation matrix for all chimpanzee cerebral cortex cells and using genes discovered in PCA of fetal neocortex single cell transcriptomes [114].These same genes had been used to infer lineage relationships in the fetal neocortex. We then generated a weighted adjacency network graph using the graph.adjacency() command and visualized cells as vertices connected to other cells via edges if the pairwise correlation between two cells was higher than 0.4. The fruchterman reingold layout was used to plot the network graph. The combined species network was con- structed in a similar way using the same genes and a correlation threshold of 0.4, and was based on FPKM quantification of alignments to each respective species' reference genome. Monocle [251] was used to establish pseudotime estimates and corroborate lineage relationships of chimpanzee cerebral cortex cells using the same genes as in the network analysis.

Human-chimpanzee consensus genome construction We re-aligned reads from each cell to a human-chimpanzee consensus genome to account for mapping bias originating from the different genome qualities of the human and chimpanzee genome. The consensus genome was generated as previously described [156]. In brief, the consensus genome was constructed based on the chained and netted pairwise alignment of human (hg38) and chimpanzee (panTro4) obtained from UCSC. Discordant sites and indels including 6 bp upstream and downsteam of the indel position were masked (replacing the base with N). STAR v2.5.1a [266] was used to map all sequences to the consensus genome requiring a minimal fraction of 85% of mapped bases per read. For quantification, HTSeq [267] v0.6.1.p1 was used applying the human GENCODE v.24 annotation. Resulting count files were combined into one master table containing all cells and genes.

Differential gene expression analysis To identify differentially expressed genes between human and chimpanzee, cells were first annotated as AP, BP or neuron based on the fetal cortex cell type with which each cell maximally correlated. After cell type assignment, SCDE (Single Cell Differential Expression) [268], a Bayesian approach for finding differentially expressed genes accounting for noise inherent to single-cell data, was used to compare the orthologous cell type between human and chimpanzee. AP or neuronal specificity was defined as one standard deviation from the mean of z-scores from SCDE of APs and Neurons (Z.x). A more stringent threshold of twice the standard deviation of the z-score was used to define differential expression between human and chimpanzee (Z.y). For the differential gene expression analysis during mitotic phases, we aimed to identify relatively homogeneous clusters of human organoid APs, chimpanzee organoid APs, endothelial cells (ECs), or iPSCs in G2M or G1 phases. We hierarchically clustered cells (Pearson correlation) using expression of genes that correlated with PC1 from PCA on human fetal cortex progenitor cells [114] and which are able to distinguish between cells in G2M and G1 phases. We selected the clusters with high or no expression and assigned them as G2/M or G1, respectively, and ignored the intermediate cells for SCDE. For the organoid APs, this assignment was consistent with an independent assignment using the method published by [269].

Immunohistofluorescence

Cerebral organoids were fixed with 1% PFA in 120 mM phosphate buffer pH 7.4 for 20 min at room temperature and subjected to cryosectioning (14 µm) and immunofluorescence as described [114]. The following primary antibodies were used: rabbit anti-PAX6 (PRB-278P; Covance), sheep anti-TBR2 (AF6166; R+D systems), rat anti-CTIP2 (ab18465; Abcam), rabbit anti-KI67 (ab15580; Abcam). The secondary antibodies, used in combination with DAPI staining, were all donkey-derived and conjugated with Alexa 488, 555 or 647 (Life Technologies). Images were acquired with a Zeiss LSM 880 Airy inverted microscope, using 10X (0.45 NA) and 20X (0.8 NA) Plan-Apochromat objectives, and analysed using Fiji. Quantifications were carried out in cortical regions of D28 and D52-54 cerebral organoids by counting, from the ventricular to the pial surface, either all PAX6 and TBR2 positive and negative nuclei stained by DAPI in 50 µm and 100 µm wide fields, respectively, or all KI67-positive cells in 100 µm wide fields. An average of 350 cells per sample were counted. Statistical significance was calculated using the Mann–Whitney U-test.

Cumulative EdU labeling

EdU was added to 52 day old cerebral organoids at a final concentration of 1 mg/ml (added from a 1 µg/ml EdU stock in PBS). The organoids were supplied with fresh medium containing EdU every six hours for up to 48 hr. Organoids were then collected in triplicates at the indicated time points (1, 2, 6, 24, 30/36, 48 hr) and processed as described above. For EdU detection, the Click-iT EdU Alexa Fluor 647 Imaging Kit (Invitrogen C10340) was used according to the manufacturer's instructions. Cell cycle parameters were determined using linear regression based on a model described previously [270].

Live imaging

Live tissue imaging was performed as described previously [247]. In short, cerebral organoids or freshly dissected developing neocortex tissue were embedded in agarose (Sigma, Germany), sectioned with a vibratome (∼200 µm, Leica, Germany), embedded in type Ia collagen (Cellmatrix, Japan), mounted in glass bottom microwell dishes (MatTek, Germany), and incubated with Hoechst 33342 (Sigma) as vital DNA dye. Tissue slices in the dish were further cultured for observation in a microscope stage incubation chamber (Pecon, Germany) kept at 37°C. iPSCs and B cells were likewise mounted in glass bottom microwell dishes previously coated for 1h with matrigel (BD Biosciecne) and poly-D-lysine (Sigma, Germany) respectively, and imaged under their respective standard culturing conditions (see above). Potential phototoxicity was stringently controlled as previously described [271].

Image analysis Images were viewed and prepared with ImageJ (http://imagej.nih.gov/ij/). Brightness and contrast of images were recorded and adjusted linearly. Spindle orientation analysis was performed as described [247]. In short, the degree values given in Figure 4 are deviations from a perfect orthogonality with the local apical surface plane, as seen from a coronal perspective (Figure 4A–F). For Figure 4G, the maximal range of orientations per every mitotic AP was calculated from the formation of a metaphase plate to anaphase onset.

Mitotic phase length determination To measure the duration of mitotic phases, the start of each different phase was defined as follows, based on morphology, dynamics and condensation of chromosomes as reported by vital DNA staining (Figures 5 and 6). Prophase: the beginning of mitotic chromosome condensation; prometaphase + metaphase: the beginning of chromosome congression and alignment; anaphase: the beginning of chromosome segregation toward the mitotic poles of the dividing cell; telophase: the beginning of chromosome decondensation after maximal chromosome condensation in late anaphase and until a level indistinguishable from interphase was achieved. The total duration of mitosis was the sum of these phases. We note that our measurements of mitotic phases are limited by the use of chromosomes as markers. Nevertheless, the use of a single fluorescence channel allowed a very high time resolution ($\sim$1.1 min) for close monitoring of key chromosomal dynamics to delimit mitotic phases. Towards distinguishing between prometaphase and metaphase, we subdivided prometaphase + metaphase into 'prometaphase', defined here as the time in which chromosomes are congressing and aligning toward the formation of a metaphase plate, and 'metaphase', defined here as the time after every chromosome has been incorporated into a tight metaphase plate at the equatorial plane of the cell, and until anaphase onset.

Statistical analysis Data were tabulated with Excel (Microsoft, Redmond, WA) and analysed with GraphPad Prism (La Jolla, CA). Statistical tests: for two groups of observations, the Mann–Whitney U-test was used. For three or more groups, the Kruskal–Wallis ANOVA in conjunction with Dunn's Multiple Comparison test for pair-wise comparisons was used. Results were interpreted as statistically significant when $p<0.05$.

Additional information

Acknowledgements

We thank the Services and Facilities of the Max Planck Institute of Molecular Cell Biology and Genetics for outstanding support, notably Jussi Helppi and his team of the Animal Facility, and Jan Peychl and his team of the Light Microscopy Facility. We thank David Andrijevic and Anne Weigert for help with maintenance and characterization of iPSC lines. We thank Marta Florio for assistance with human tissue dissection. We thank Andrea Musacchio and members of the Huttner, Treutlein and Pääbo labs for helpful discussions. We thank Fred Gage and Rick Livesey for kindly donating the PR818-5 iPSC line. SK was supported by a PhD fellowship of the Boehringer Ingelheim Fonds. SP was supported by the Paul G. Allen Family Foundation. WBH was supported by grants from the Deutsche Forschungsgemeinschaft (DFG, SFB 655, A2) and the European Research Council (ERC, 250197), by the DFG-funded Center for Regenerative Therapies Dresden, and by the Fonds der Chemischen Industrie. SP, BT and WBH were supported by the Max Planck Society.

Author contributions

FM-B, Conceived the study, Designed the experiments, Performed and analysed live imaging experiments, Wrote the paper; FB, Conceived the study, Designed the experiments, Grew cerebral organoids, Performed and analysed organoid immunohistochemistry and cumulative EdU labelling, Wrote the paper; SK, Conceived the study, Designed the experiments, Grew cerebral organoids, Performed single-cell RNA-seq experiments, Analysed single-cell RNA-seq data, Wrote the paper; JGC, Conceived the study, Designed the experiments, Performed single-cell RNA-seq experiments, Analysed single-cell RNA-seq data, Wrote the paper; BVe, Analysed single-cell RNA-seq data, Provided information relevant for the interpretation of the data; KK, BVo, KO, TM, Prepared chimpanzee iPSC line Sandra A and orangutan iPSC line Toba, Provided information relevant for the interpretation of the data; ZH, Constructed human-chimpanzee consensus genome ; RL, Provided human fetal tissue, Provided information relevant for the interpretation of the data; SP, WBH, Conceived the study, Designed the experiments, Provided intellectual guidance in the interpretation of the data, Wrote the paper; BT, Conceived the study, Designed the experiments, Analysed single-cell RNA-seq data, Provided intellectual guidance in the interpretation of the data, Wrote the paper.

Ethics

Human subjects: Human fetal brain tissue (11-13 weeks post conception (wpc)) was obtained with informed written maternal consent followed by elective pregnancy termination. Research involving human tissue was approved by the Ethical Review Committee of the Universitätsklinikum Carl Gustav Carus of the Technische Universitaet Dresden. In addition, research was approved by the Institutional Review Board of the Max Planck Institute of Molecular Cell Biology and Genetics. Animal experimentation: Mice were kept pathogen-free at the Biomedical Services Facility of the MPI-CBG. All experiments using mice were performed according to the German Animal Welfare Legislation. In addition, research was approved by the Institutional Review Board of the Max Planck Institute of Molecular Cell Biology and Genetics.

Additional files

Major datasets
The following dataset was generated:

Kanton S, Camp G, Treulein B, 2016
Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development
https://www.ncbi.nlm.nih.gov/geo/query/acc. cgi?acc=GSE86207
Publicly available at NCBI Gene Expression Omnibus (accession no: GSE86207)

Figure supplements

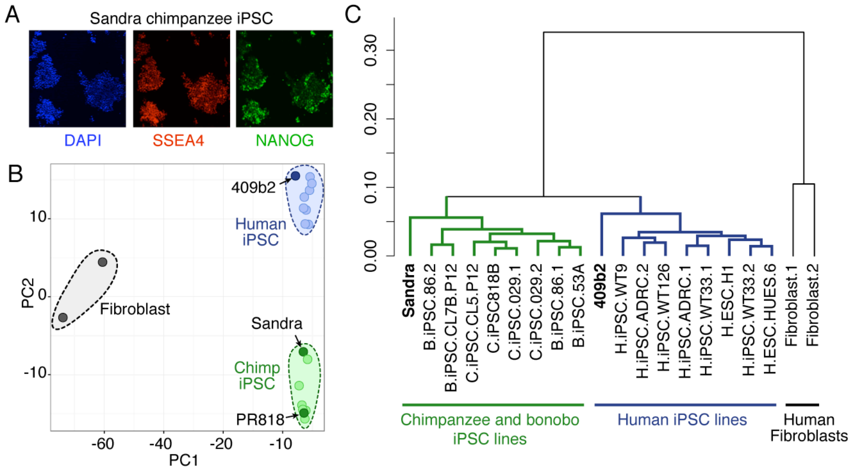


Figure 1-figure supplement 1: Characterization of chimpanzee iPSCs. (**A**) Chimpanzee iPSC line Sandra A stained for pluripotency markers SSEA5 (red) and NANOG (Green). Nuclei are stained with DAPI. (**B**) PCA on bulk RNA-seq data from human iPSCs, chimpanzee and bonobo iPSCs, and human fibroblasts was used to describe the variation between cell types. RNA-seq data on chimp iPSC line Sandra A and human iPSC line 409b2 was generated in this study. Data from the other human, chimpanzee and bonobo, and fibroblast lines were previously published [82, 272]. (**C**) Dendrogram showing hierarchical clustering of human IPSC, chimpanzee and bonobo IPSC, and human fibroblast lines based on the Pearson correlation of the expression of 12,221 genes.

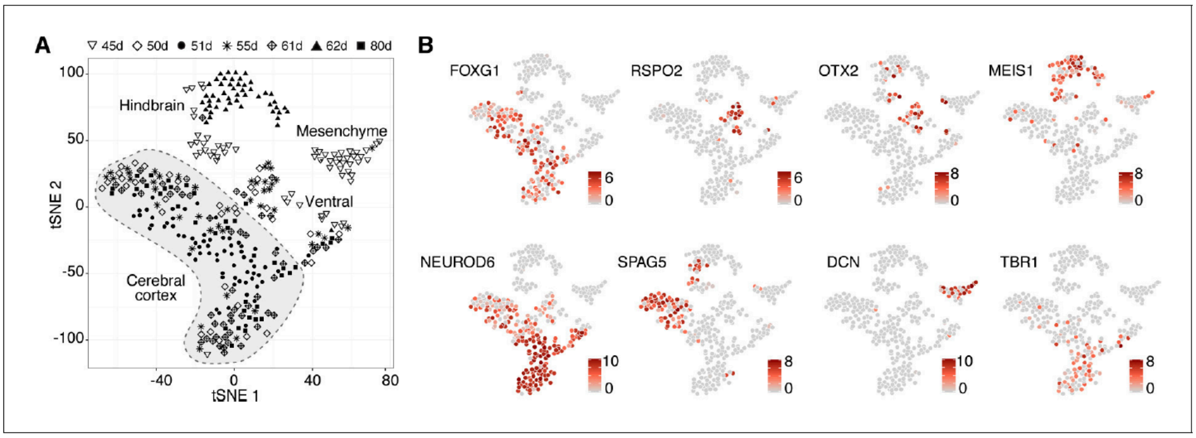


Figure 1-figure supplement 2: Deconstructing cell type composition in chimpanzee cerebral organoids using single-cell RNA-seq. (**A**) scRNA-seq was performed on chimpanzee organoids dissociated at 45, 50, 51, 55, 62, and 80 days (d) after embryoid body (EB) culture. PCA and unbiased clustering using tSNE reveals cell populations from hindbrain, midbrain, mesenchyme, and cerebral cortex (shaded in grey) within organoids. Different symbols indicate different experiments. (**B**) Marker genes are shown for each cluster with cells coloured based on gene expression level. Cerebral cortex cells have high expression of *FOXG1* and *NEUROD6*, and low expression of *OTX2* and *RSPO2*. Progenitors express marker *SPAG5*. Cells are coloured based on expression level.

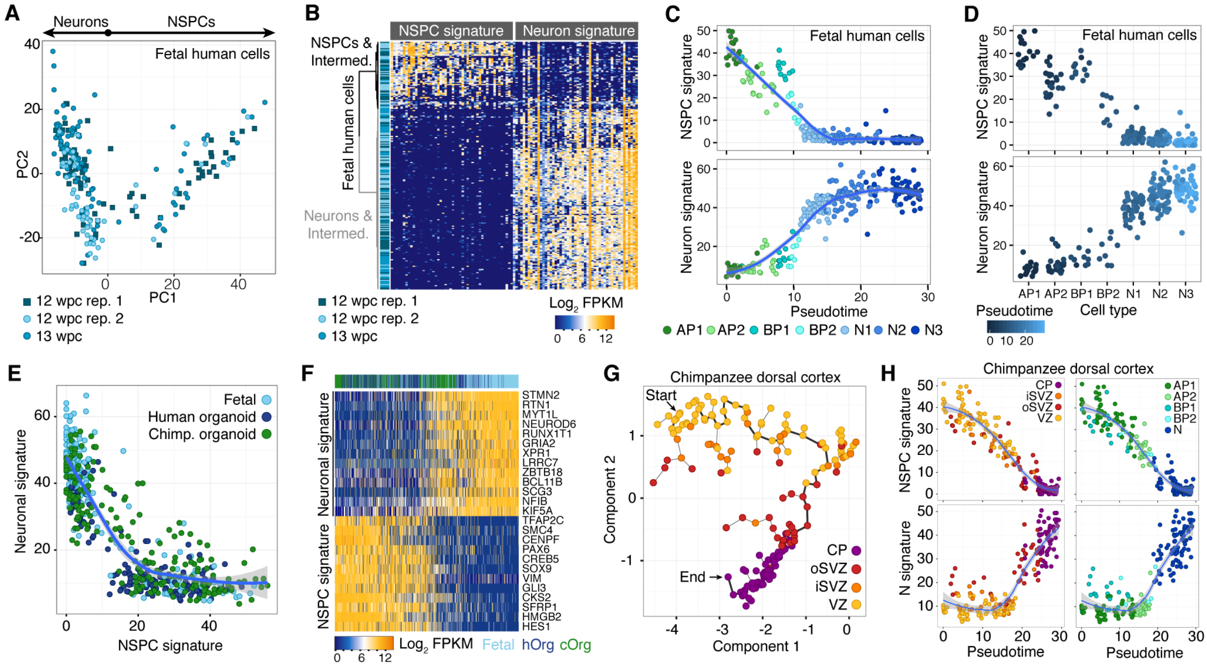


Figure 1-figure supplement 3: Fetal human progenitor and neuronal neocortical signatures are recapitulated in chimpanzee cerebral organoids. (**A**) PCA of human fetal neocortex was used to identify genes describing cortical cell populations. Each dot represents a cell that is color-coded in shades of blue representing three different experiments. The genes correlating and anticorrelating with PC1 were used to define the NSPC and neuron signature, respectively. (**B**) Hierarchical clustering and heatmap visualization showing the expression of genes that have highest correlation (NSPC signature) and anti-correlation (Neuron signature) with PC1. Cells are shown in rows, genes in columns. (**C,D**) Fetal cortical cells were classified as APs in G2-M (AP1), APs in G1-S (AP2), BPs in G2-M (BP1), BPs in G1-S (BP2), or migrating (N1, N2) and cortical plate (N3) neurons. Each cell was scored for the NSPC (top) or neuron (bottom) signature and plotted in the order of pseudotemporal point on the neurogenic lineage (**C**) or plotted for each cell type (**D**). (**E**) Scatterplot showing NSPC and neuronal signature scores for each human fetal, human organoid and chimpanzee organoid cortical cell. The signatures were derived from PCA of fetal cerebral cortex single-cell transcriptomes. (**F**) Heatmap showing gene expression of top NSPC and neuron signature genes across human fetal, human organoid (hOrg), and chimpanzee organoid (cOrg) cells. (**G**) Monocle reveals a NSPC-to-neuron lineage in the chimpanzee organoid that correlates with the zones of the developing fetal primate neocortex. Cells (circles, coloured by maximum correlation with cortical zones; CP, cortical plate) are arranged in the 2-D independent component space based on genes identified using PCA. The minimal spanning tree (grey lines) connects cells, with the black line indicating the longest path. (**H**) Each chimpanzee cerebral organoid cortical cells scored for the NSPC (top) or neuron (N, bottom) signature and plotted in the order of pseudotemporal position on the neurogenic lineage. Cells are coloured by maximum correlation with cortical zones (left) or cell type (right).

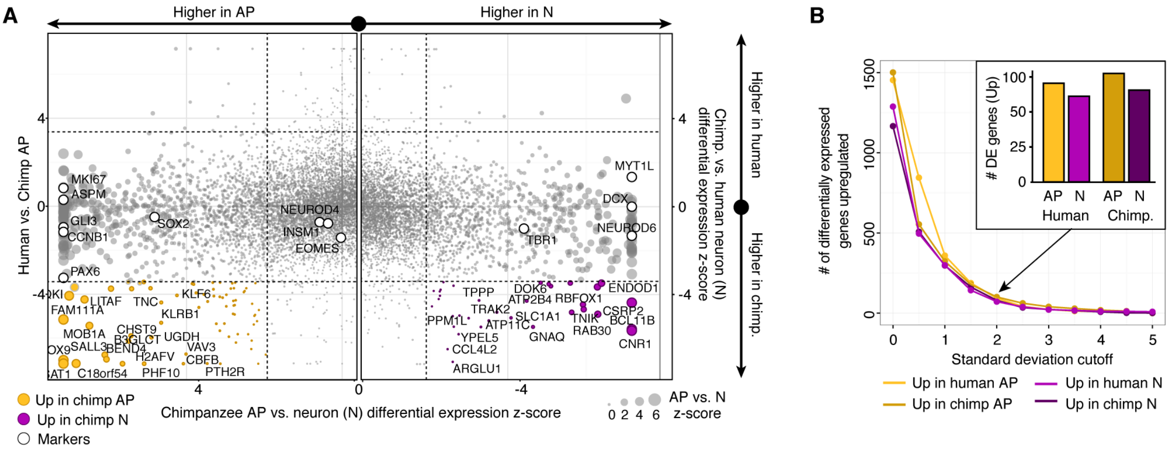


Figure 3-figure supplement 1: Differential expression analysis between chimpanzee and human cerebral cortex cells from the chimpanzee perspective. (**A**) Scatterplots showing z-scored significance estimates from single-cell differential expression (SCDE) analysis based on Bayesian probabilistic models. Reads from human and chimpanzee were mapped to a consensus genome, and human gene annotations were used for expression counting. The x-axis represents SCDE between chimpanzee organoid APs vs. chimpanzee organoid neurons (N). The y-axes on the left and right plots represents SCDE between human and chimpanzee APs and neurons, respectively. Genes coloured as white circles represent marker genes from Figure 1 and are generally not differentially expressed between human and chimpanzee, but do vary between chimpanzee APs and neurons, validating the SCDE analysis. Yellow and purple circles represent genes upregulated specifically in chimpanzee APs and neurons, respectively. Circles are sized based on differential expression between chimpanzee APs and neurons. (**B**) Plot showing the number of differentially expressed genes between human and chimpanzee cells as a function of standard deviations above the mean z-score from the Bayesian differential gene expression analysis.

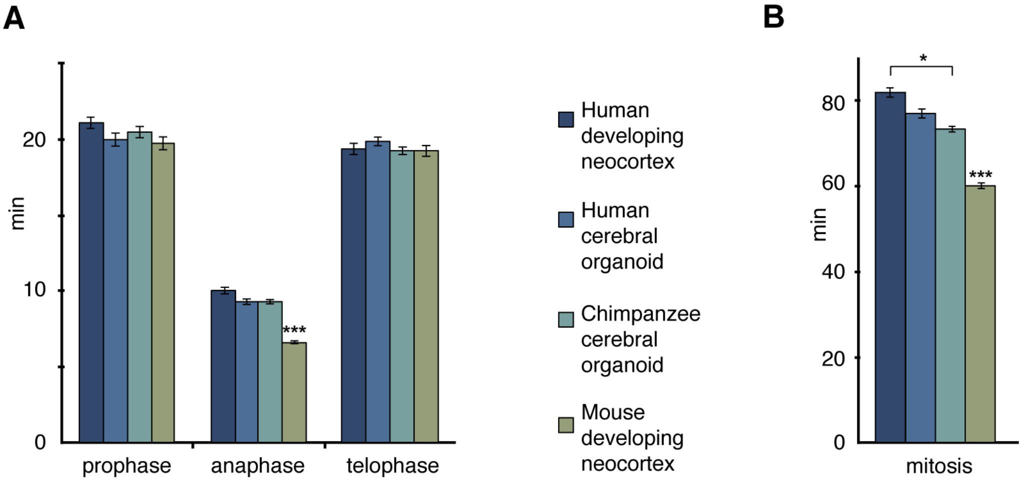


Figure 5-figure supplement 1: The length of the mitotic phases other than prometaphase-metaphase is similar between human and chimpanzee APs. Length of prophase, anaphase and telophase (**A**), and of total mitosis (**B**, sum of all mitotic phases described here and in Figure 5) between APs of human developing neocortex, humanand chimpanzee cerebral organoids and mouse developing neocortex, determined from the experiments described in Figure 5. Data are the mean ± SEM of ≥60 APs from ≥4 independent experiments each. *p<0.05; ***p <0.001.

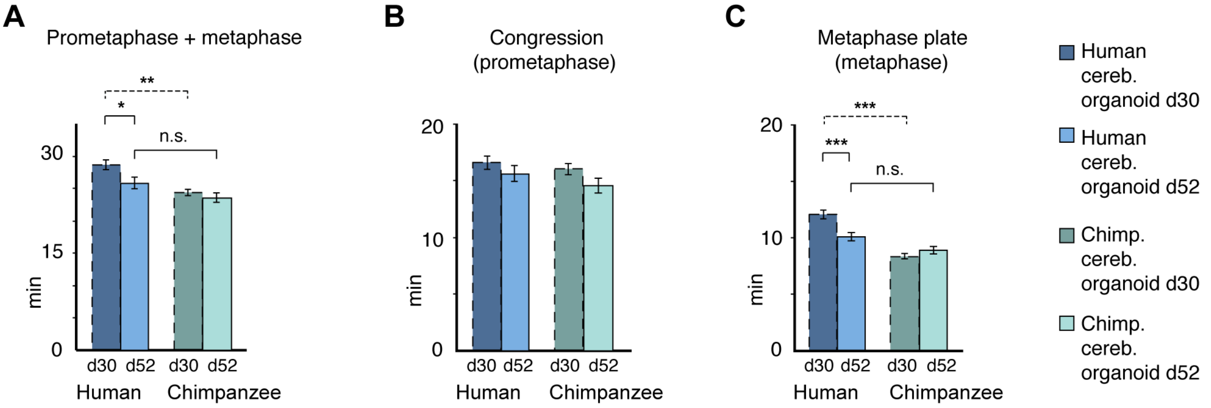


Figure 5-figure supplement 2: Differences in prometaphase-metaphase length between APs of D30 and D52 human and chimpanzee cerebral organoids. Mitotic phase measurements similar to those in Figure 5E–G, but for APs in D52 organoids. Time between the start of chromosome congression and anaphase onset (referred to as 'prometaphase + metaphase') (**A**), between the start of chromosome congression and the formation of a metaphase plate (referred to as 'prometaphase') (**B**), and between the formation of a metaphase plate and anaphase onset (referred to as 'metaphase') (**C**). Data include APs from organoids from the human iPSC line SC102A-1 and chimpanzee iPSC line Sandra A, and are the mean ± SEM of 30 APs from 2 independent experiments each. For comparison, the relevant data for human and chimpanzee D30 cerebral organoid APs from Figure 5 are shown (dashed lines). *$p<0.05$; **$p<0.01$; ***$p<0.001$; n.s., not significant.

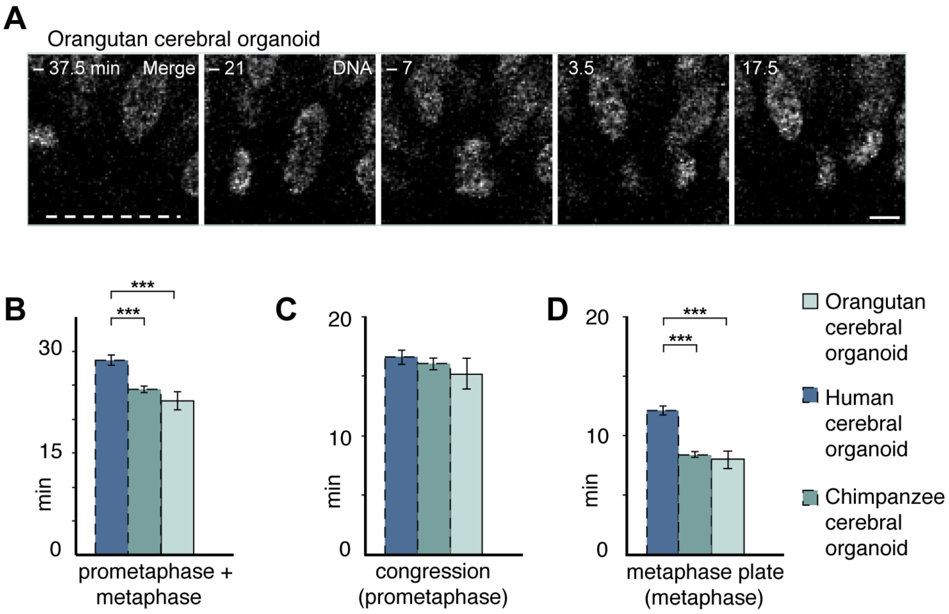


Figure 5-figure supplement 3: Prometaphase-metaphase in orangutan organoid APs is similar to chimpanzee organoid APs. Live tissue imaging of mitotic phases, as reported by chromosomes, in organotypic slice culture of orangutan cerebral organoid. 0 min is anaphase onset. Time-lapse is ~ 3.5 min. (**A**) AP in a slice of orangutan D30 cerebral organoid (Toba). The time indicated on each image is when that image was taken, relative to anaphase onset (0 min). White dashed line, ventricular surface. Scale bar, 5 µm. (B–D) Time between the start of chromosome congression and anaphase onset (referred to as 'prometaphase + metaphase') (**B**), between the start of chromosome congression and the formation of a metaphase plate (referred to as 'prometaphase') (**C**), and between the formation of a metaphase plate and anaphase onset (referred to as 'metaphase') (**D**). For comparison, the relevant mitotic phase lengths of human and chimpanzee cerebral organoid APs from Figure 5 are shown (columns with dashed line). Data for orangutan are the mean ± SEM of 16 cells from 2 different cortex regions of an organoid. ***$p<0.001$.

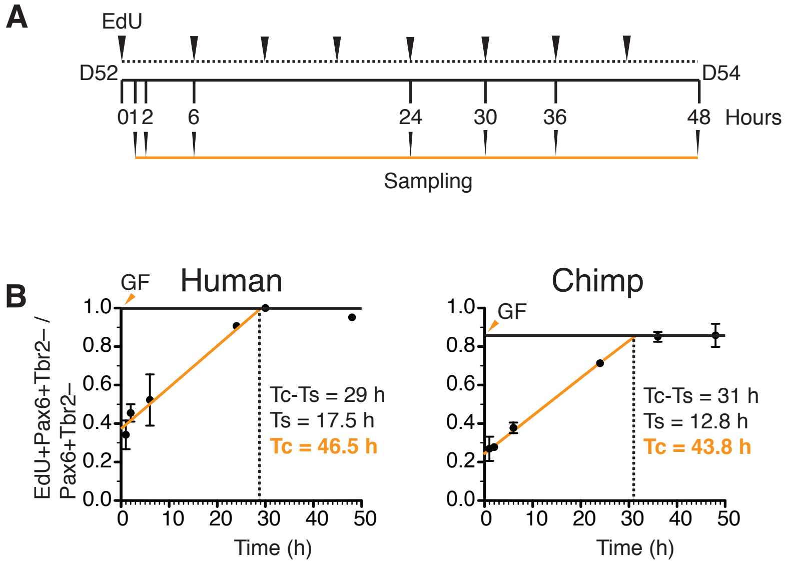


Figure 5-figure supplement 4: Determination of cell cycle parameters of human and chimpanzee organoid APs using cumulative EdU labeling. (**A**) Schematic representation of the cumulative EdU labeling experiment. (**B**) Linear regression curves of human ($r^2 = 0.984$) and chimpanzee ($r^2 = 0.998$) PAX6+TBR2– cells after cumulative EdU labeling over 48 hr. The total cell cycle length minus the S-phase length (Tc-Ts) was obtained directly from the graph (vertical dashed lines). The S-phase length (Ts) was determined from the x-axis and y-axis intercepts of the linear regression curves, and the total cell cycle length (Tc) was calculated therefrom. The growth fraction (GF, solid horizontal line) is indicated by orange arrowheads.

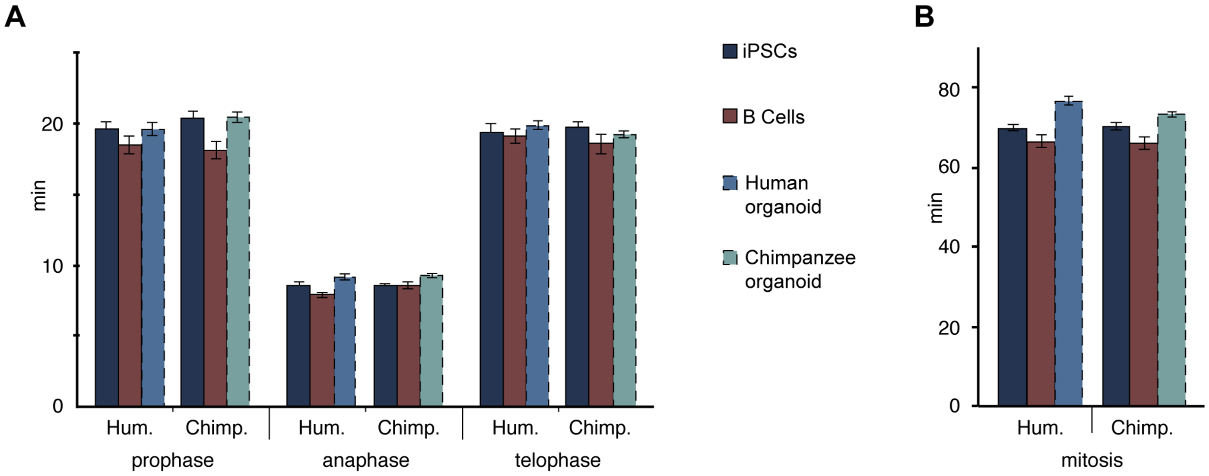


Figure 6-figure supplement 1: The length of prophase, anaphase and telophase is similar in human and chimpanzee iPSCs, B cells and organoid APs. Length of prophase, anaphase and telophase (**A**), and of total mitosis (**B**, sum of all mitotic phases described here and in Figure 6) in human and chimpanzee iPSCs and B cells, determined from the experiments described in Figure 6. For comparison, the relevant mitotic phase lengths of human and chimpanzee cerebral organoid APs from Figure 5–figure supplement 1 are shown (columns with dashed line). Data are the mean ± SEM of ≥30 cells from ≥3 independent experiments each.

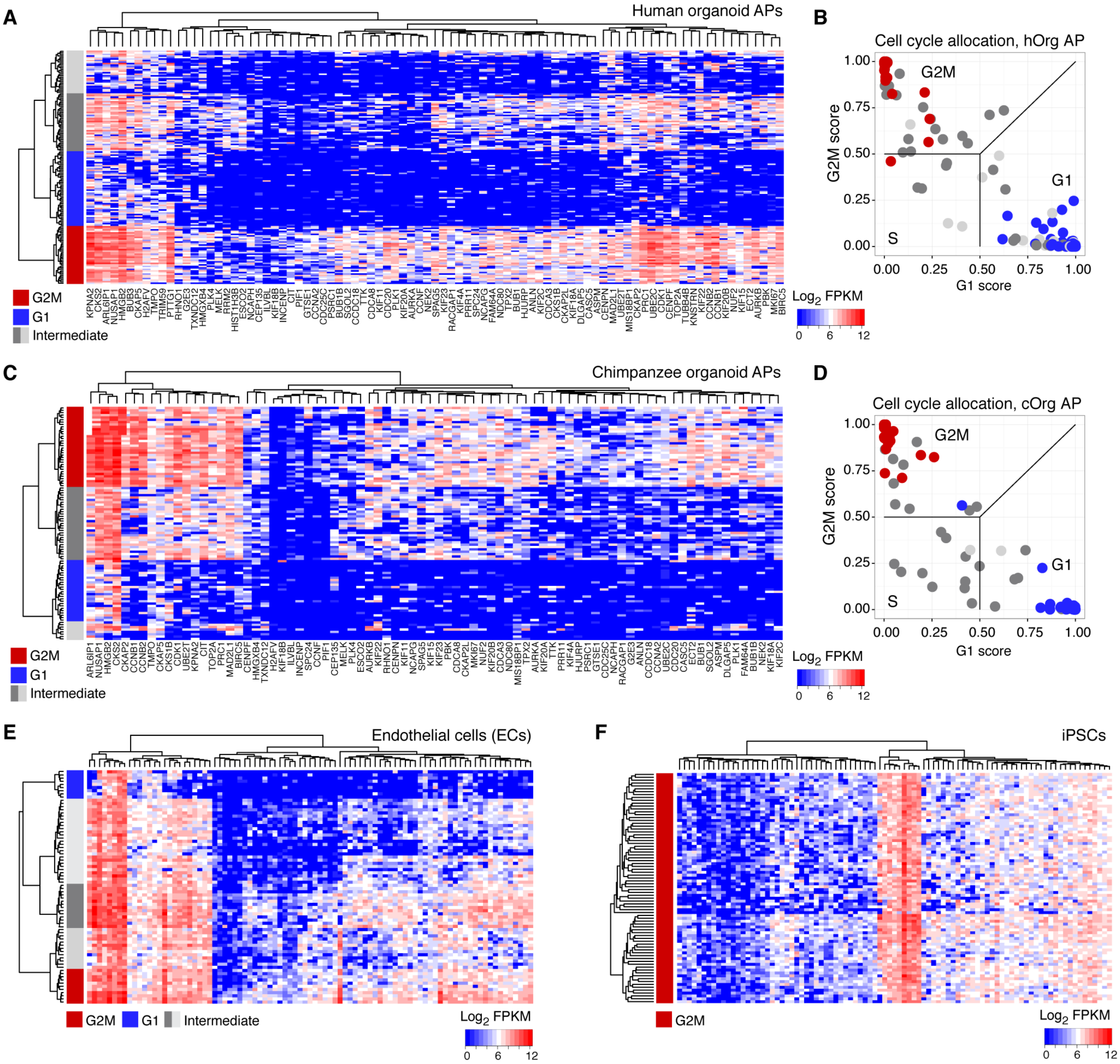


Figure 8-figure supplement 1: Cell cycle assignment for differential gene expression analysis. (**A**) Hierarchical clustering was used to identify human organoid APs that most strongly expressed genes enriched in G2M phase of the cell cycle (red). The genes were identified from PCA on fetal cortex progenitor cells (top 100 correlating genes) [114]. The cluster with weakest expression of these G2M associated genes was assigned as G1 phase (blue). Intermediate cells (grey) were discarded from differential gene expression analysis. (**B**) A previously published method was used to computationally assign cell-cycle stage based on a machine-learning approach [269]. This method was generally consistent with our assignment based on the hierarchical clustering presented in panel A. (**C**–**F**) The same approach was used to identify the chimpanzee organoid APs, endothelial cells (ECs), and iPSCs that most highly express G2M markers. Note that all iPSCs analyzed highly expressed most of the G2M marker genes.

Source data

Source data can be found in the online version of the article (doi: 10.7554/eLife.18683)

Figure 1—source data 1

Processed single-cell RNA-seq data for chimpanzee cells.

*.txt file containing processed chimpanzee single-cell RNA-seq data (344 single cells) in log2(FPKM) with genes in columns and cells in rows. The first 7 columns contain metadata for each cell: cortex: assignment of cell to cortex (1) or to other regions within organoid(0); tSNE_1: tSNE1 loading for each cell; tSNE_2: tSNE2 loading for each cell; PC1: PC1 loading for each cell; PC2: PC2 loading for each cell; species: species of origin for each cell; cell_id: unique ID for each cell, with information about the experiment and the age of the organoid of origin for each cell.

https://doi.org/10.7554/eLife.18683.004

Figure 1—source data 2

Genes describing cell populations in the chimpanzee organoids.

List of genes identified by PCA on all chimpanzee organoid single-cell transcriptomes as being most informative for defining cell populations.

https://doi.org/10.7554/eLife.18683.005

Figure 3—source data 1

Processed single-cell RNA-seq data for human cells.

*.txt file containing processed human single-cell RNA-seq data (207 single cells) in log2(FPKM) with metadata in first 4 columns for each cell: cell_id: unique ID for each cell; experiment: the experiment during which each cell was isolated; species: species of origin for each cell; cortex: assignment of cell to cortex (1) or to other regions within organoid (0).

https://doi.org/10.7554/eLife.18683.011

Figure 3—source data 2

Results of differential gene expression analyses.

Excel file (*.xlsx) with multiple sheets containing results of all differential expression analyses presented in the manuscript as well as GO enrichment analysis for the differentially expressed (DE) genes: Sheet 1: Genes specific to APs, not DE between chimpanzee and human; Sheet 2: GO enrichment analysis for genes of sheet 1; Sheet 3: Genes specific to Neurons, not DE between chimpanzee and human; Sheet 4: GO enrichment analysis for genes of sheet 3; Sheet 5: Genes specific to APs and upregulated to human compared to chimpanzee; Sheet 6: GO enrichment analysis for genes of sheet 6; Sheet 7: Genes specific to Neurons and upregulated to human compared to chimpanzee; Sheet 8: GO enrichment analysis for genes of sheet 7; Sheet 9: Genes specific to APs and upregulated to chimpanzee compared to human; Sheet

10: GO enrichment analysis for genes of sheet 6; Sheet 11: Genes specific to Neurons and upregulated to chimpanzee compared to human; Sheet 12: GO enrichment analysis for genes of sheet 11; Sheet 13: GO enrichment data used to generate Figure 3F.
https://doi.org/10.7554/eLife.18683.012

Figure 5—source data 1
Durations of all mitotic phases.
Numerical values in minutes for the duration of all mitotic phases ±SEM used in the graphs in Figures 5, 6 and 7, in Figure 5—figure supplement 1, 2 and 3, and in Figure 6—figure supplement 1.
https://doi.org/10.7554/eLife.18683.016

Chapter 2

Organoid single-cell genomic atlas uncovers human-specific features of brain development

Sabina Kanton, Michael James Boyle, Zhisong He, Malgorzata Santel, Anne Weigert, Fatima Sanchis Calleja, Patricia Guijarro, Leila Sidow, Jonas S. Fleck, Dingding Han, Zhengzong Qian, Michael Heide, Wieland B. Huttner, Philipp Khaitovich, Svante Pääbo, Barbara Treutlein and J. Gray Camp

Nature, Vol. 574, Issue 7778 (2019)

Organoid single-cell genomic atlas uncovers human-specific features of brain development

Sabina Kanton[1,7], Michael James Boyle[1,7], Zhisong He[1,2,7]*, Malgorzata Santel[1], Anne Weigert[1], Fatima Sanchis Calleja[1,2], Patricia Guijarro[3], Leila Sidow[1], Jonas Fleck[2], Dingding Han[3], Zhengzong Qian[3], Michael Heide[4], Wieland B. Huttner[4], Philipp Khaitovich[1,3,5], Svante Pääbo[1], Barbara Treutlein[1,2]* and J. Gray Camp[1,6]*

Affiliations:

[1]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

[2]Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland.

[3]CAS Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai, China.

[4]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany.

[5]Center for Neurobiology and Brain Restoration, Skolkovo Institute of Science and Technology, Moscow, Russia.

[6]Institute of Molecular and Clinical Ophthalmology, Basel, Switzerland.

Correspondence to: zhisong.he@bsse.ethz.ch
barbara.treutlein@bsse.ethz.ch
grayson.camp@iob.ch

[7]These authors contributed equally to this work

Abstract

The human brain has undergone substantial change since humans diverged from chimpanzees and the other great apes [223, 273]. However, the genetic and developmental programs underlying this divergence are not fully understood. Here we have analysed stem cell-derived cerebral organoids using single-cell transcriptomics and accessible chromatin profiling to investigate gene-regulatory changes that are specific to humans. We first analysed cell composition and reconstructed differentiation trajectories over the entire course of human cerebral organoid development from pluripotency, through neuroectoderm and neuroepithelial stages, followed by divergence into neuronal fates within the dorsal and ventral forebrain, midbrain and hindbrain regions. Brain-region composition varied in organoids from different iPSC lines, but regional gene expression patterns remained largely reproducible across individuals. We analysed chimpanzee and macaque cerebral organoids and found that human neuronal development occurs at a slower pace relative to the other two primates. Using pseudotemporal alignment of differentiation paths, we found that human-specific gene expression resolved to distinct cell states along progenitor-to-neuron lineages in the cortex. Chromatin accessibility was dynamic during cortex development, and we identified divergence in accessibility between human and chimpanzee that correlated with human-specific gene expression and genetic change. Finally, we mapped human-specific expression in adult prefrontal cortex using single-nucleus RNA sequencing analysis and identified developmental differences that persist into adulthood, as well as cell-state-specific changes that occur exclusively in the adult brain. Our data provide a temporal cell atlas of great ape forebrain development, and illuminate dynamic gene-regulatory features that are unique to humans.

Main text

Bulk genomic measurements in primary brain tissue from adult humans, chimpanzees and other apes [46, 55, 141, 144], as well as from developing rhesus macaque [140, 151], have identified molecular features that appear to be specific to the human brain. These studies have largely been limited by ensemble averaging and it has been difficult to perform similar experiments in developing great apes owing to the lack of available tissue. Cerebral organoids [92] grown from great ape induced pluripotent stem cells (iPSCs) [82] offer the potential to study the evolution of human brain development in controlled culture environments. Previous single-cell sequencing studies have shown that human organoids can recapitulate many aspects of in vivo cortex development [114, 116, 117, 128], and comparisons with other primate organoids can be used to identify human-specific gene expression [20, 222]. However, the gene-regulatory mechanisms that distinguish humans from the other primates are unclear. Here we use single-cell RNA sequencing (scRNA-seq) together with accessible chromatin profiling to understand human cerebral organoid development and explore how human cortical gene expression programs have diverged from chimpanzee. We further analysed adult prefrontal cortex tissue using single-nucleus RNA-seq (snRNA-seq) to reveal the potential and limits of cerebral organoids to study human-specific expression patterns.

We first used scRNA-seq (10x Genomics) to profile cell composition across a time course of human organoid development from pluripotency to four months using embryonic stem cells (H9) and an iPSC (409b2) line (Fig. 1a, Extended Data Fig. 1; 43,498 cells). We constructed a force-directed *k*-nearest neighbour graph [274] based on pseudocells to visualize the temporal progression (Fig. 1b). We tracked a progression through stem cell states during the first 15 days, and by 1 month, cells diversified into neural progenitor cells (NPCs) of multiple brain regions including the forebrain (dorsal and ventral telencephalon and diencephalon), midbrain (mesencephalon), hindbrain (rhombencephalon) and retina. By two months, excitatory and inhibitory neuronal fates had differentiated, and by four months, astrocytes had emerged. These observations were based on inspection of marker genes, comparisons to transcriptomes from regions and cells from the developing human brain [76, 219], and analysis of spliced and unspliced transcripts [275] (Fig. 1c,d, Extended Data Fig. 2).

We next analysed the reproducibility of gene expression patterns across pluripotent stem cell (PSC) lines from different human individuals in two-month-old organoids (Fig. 1e, Extended Data Fig. 3). We identified cells on the neuronal lineage (49,153 cells), classified distinct cell types from different brain regions, analysed differentiation trajectories and ordered ventral and dorsal telencephalic cells along pseudotimes (Fig. 1e–g, Extended Data Fig. 3). Consistent with previous studies [113, 128], we found that each inducible PSC (iPSC) line contributed cells to multiple differentiation trajectories; however, the proportions of cells in each trajectory varied across organoid and line (Fig. 1f, Extended Data Fig. 3). Nevertheless, gene-expression patterns of each trajectory across the lines were highly correlated and cells representing each region clustered together (Fig. 1h, i, Extended Data Fig. 3). These data provide a temporally and pseudotemporally resolved gene-expression atlas of the earliest stages of human brain development and provide a baseline for identifying human-specific gene expression.

We next generated a gene-expression atlas of chimpanzee organoid development to dissect features that differ from that of the human (Fig. 2a, Extended Data Fig. 4; 36,884 cells). We identified dorsal and ventral telencephalon trajectories, as well as rhombencephalic cell populations in chimpanzee organoids (Fig. 2b). We used time warping to align human and chimpanzee cortical pseudotimes and observed that the latest pseudotime points in chimpanzee failed to map to a human counterpart (Fig. 2c). We found that chimpanzee neurons express higher levels of neuron maturation-related genes (Fig. 2d) and confirmed differences in maturation using scRNA-seq data from additional human and chimpanzee individuals [20, 222], as well as from macaque (Extended Data Figs. 5, 6). In addition, more astrocytes were observed in chimpanzee organoids compared to humans at four months (Fig. 2e). We observed more

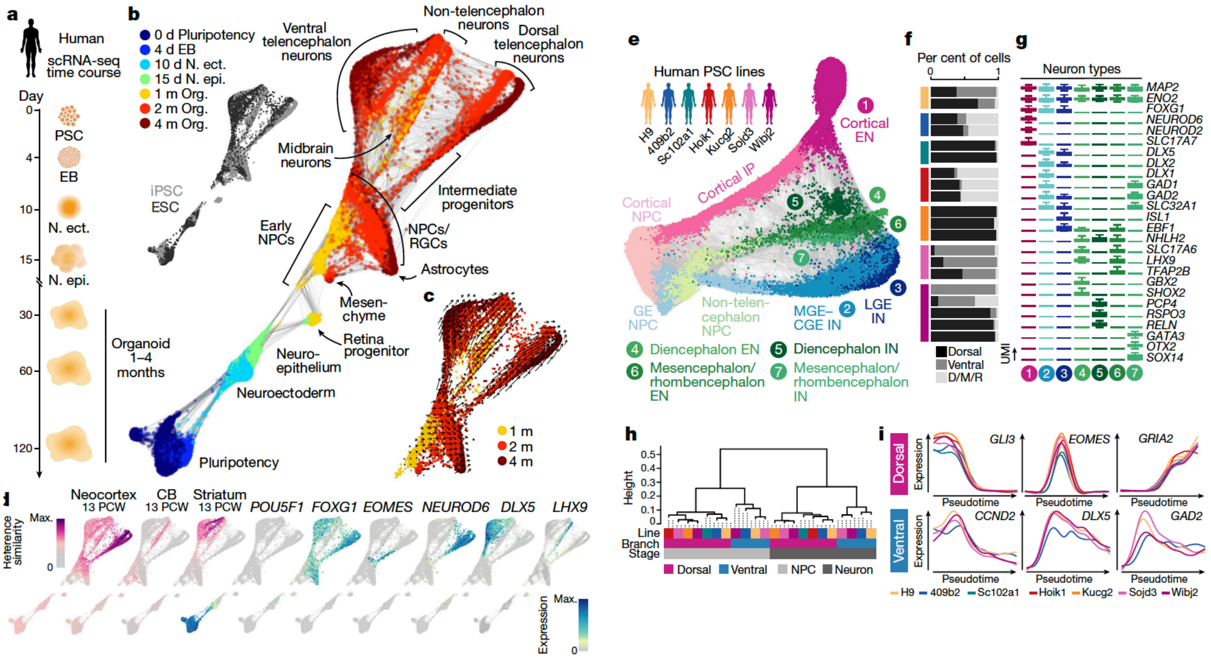


Fig. 1: Reconstructing human cerebral organoid differentiation from pluripotency. **a**, scRNA-seq was performed on iPSC- and embryonic stem cell (ESC)-derived cells (43,498) at different time points during cerebral organoid differentiation from pluripotency. EB, embryoid body; N. ect., neuroectoderm; N. epi., neuroepithelium. **b**, All time points were combined, pseudocells (11,427) were constructed and the differentiation trajectory was reconstructed using SPRING [274]. Pseudocells are coloured by time point (main image) or cell line (top left). Org., organoid; RGC, radial glial cell. **c**, RNA velocity analysis [275] supports differentiation of NPCs into distinct regions of the developing human brain. **d**, Left, SPRING plot coloured (magenta) by reference similarity spectrum (RSS) to bulk RNA-seq data generated from diverse brain regions at different time points (Allen Brain Atlas) [76]. Shown are the tissues and time points with maximum correlation. CB, cerebellum. Right, SPRING plot coloured (cyan) by marker gene expression. PCW, post-conception weeks. **e**, SPRING reconstruction based on the RSS of organoid scRNAseq data from six iPSC lines and one ESC (H9) line (49,153 cells), with clusters coloured by cell type. CGE, caudal ganglionic eminence; MGE, medial ganglionic eminence; LGE, lateral ganglionic eminence; IN, inhibitory neuron; EN, excitatory neuron. IP, intermediate progenitors; GE, ganglionic eminence. **f**, Proportion of cells per organoid that are in the dorsal or ventral telencephalon or diencephalon, mesencephalon and rhombencephalon (D/M/R) neuronal branches. **g**, Box plots (outliers removed) showing relative expression of marker genes for major neuron populations that emerge in the human cerebral organoids. Boxes represent interquartile range (IQR) and whiskers represent minimum and maximum with outliers removed. UMI, unique molecular identifier. **h**, Dendrogram based on pairwise correlations between cells from different lines, branches or stages based on pseudotime-dependent gene expression patterns. i, Pseudotemporal expression patterns of neuronal differentiation markers for the dorsal (top) and ventral (bottom) telencephalon trajectories for each line.

pronounced upper- and deep-layer cortical neuronspecification in chimpanzee and macaque organoids relative to human organoids at the same time point on the basis of expression, but this was not consistent across lines, organoids and protocols (Extended Data Fig. 6). Our data suggest that delayed maturation of the human brain [55, 57, 276] is observed in organoids and can be traced back to early stages of development, consistent with previous reports comparing human and macaque brain development in vivo [140] and in in vitro 2D cultures [59, 89].

We next aimed to detect human-specific changes in gene expression in the developing cortex. We first inspected the expression of duplicated or rearranged genes (Supplementary Table 7) and detected 23 of 24 genes included in the annotation, 4 of them specific to G2M phase progenitors of the telencephalon (*ARHGAP11B*, *FAM72B*, *FAM72C* and *FAM72D*) [61]. We next aligned all human, chimpanzee and macaque reads to a consensus genome and aligned dorsal telencephalic trajectories between the species (Fig. 2f, Extended Data Fig. 7). We then searched for genes that were differentially expressed specifically on the human branch, and found that most of the observed human-specific deviations from chimpanzee and macaque were expression gains (Fig. 2g); this was also observed for chimpanzee-specific changes (Extended Data Fig. 7). We propose that this is because it is more deleterious to lose a highly conserved gene expression pattern than it is to gain the expression of a new gene. Of the 98 identified differentially expressed genes, 96 clustered into 7 pseudotemporal patterns (Fig. 2h–j), with clusters 1, 2 and 3 specifically enriched in human radial glia, intermediate progenitors and neurons, respectively (Fig. 2h). Genes with human-specific expression gain are associated with diverse biological processes including proliferation of radial glia, neuron migration and neurite formation, and are localized to different clusters of maturing neurons (Fig. 2i, j, Extended Data Fig. 7). When comparing our results to previously published data on human and great ape organoids and fetal brains [20, 222], we find strong overlap between datasets (Extended Data Fig. 7). Using cells of ventral telencephalon identity, we find 92 genes differentially expressed between human and chimp - 17% of them distinct from those differentially expressed in the cortex (Extended Data Fig. 7). In sum, this analysis identifies human-specific gene expression changes that may be specific to certain cell states within the developing human forebrain.

To identify potential regulatory mechanisms, we performed bulk and single-cell accessible chromatin profiling (scATAC-seq; Fluidigm C1) along the organoid-differentiation time course in human and chimpanzee, including that of microdissected cortical regions (Fig. 3a–c, Extended data Fig. 8). We analysed cell heterogeneity, ordered cells in pseudotime and dynamically monitored transcription factor binding motifs and chromatin accessibility over the differentiation path (Fig. 3d, Extended Data Fig. 8). We searched for differential accessibility between human and chimpanzee cortical NPCs and neurons, and found that 7% and 9% of accessible peaks showed increased and decreased accessibility in humans, respectively (Fig. 3e), with many differential accessibility peaks being specific to either NPCs (53.8%) or neurons

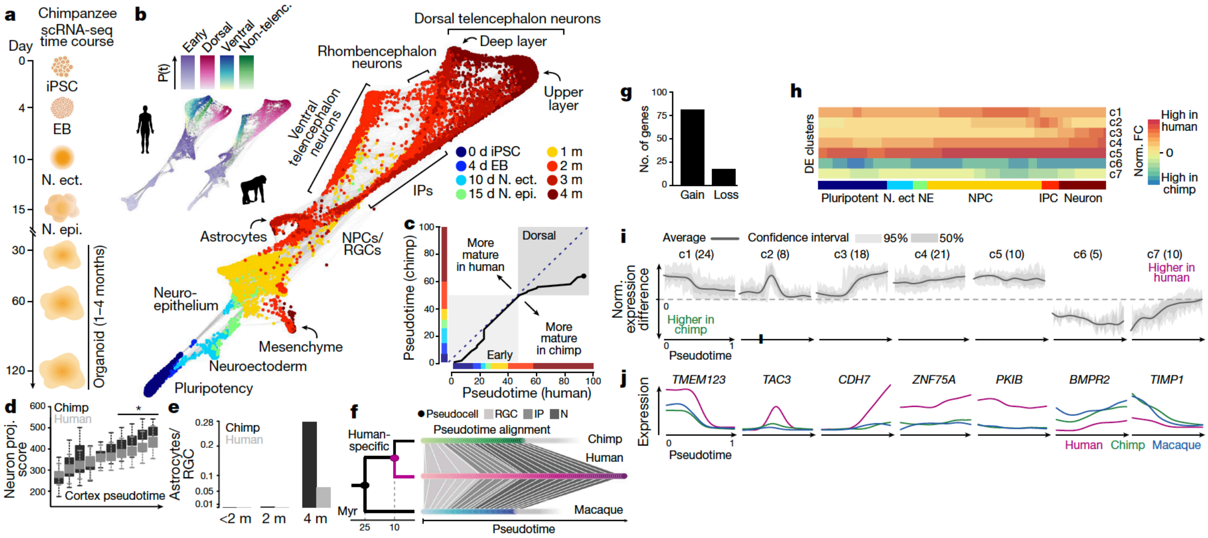


Fig. 2: Comparing human and chimpanzee organoids reveals human-specific features of cortex development. **a**, scRNA-seq was performed on chimpanzee iPSC-derived cells at different time points during cerebral organoid differentiation from pluripotency. **b**, The differentiation trajectory was reconstructed using SPRING. Pseudocells (9,647) coloured on the basis of time point (main image) or by trajectory pseudotimes (top left, alongside human data). **c**, Alignment of human and chimpanzee pseudotimes after combining pseudocells from the early stages and the dorsal forebrain lineage. **d**, Box plots (outliers removed) showing cumulative expression of neuron projection-related genes in human and chimpanzee along unaligned cortical pseudotimes. Boxes represent IQR and whiskers represent minimum and maximum with outliers removed. **e**, Astrocytes identified by scRNA-seq in organoids at different time points, normalized by the corresponding number of radial glia. **f**, Schematic of pseudotime alignment from cortical NPC to deeper layer neurons in human, chimpanzee and macaque. **g**, Number of differentially expressed genes in human versus chimpanzee and macaque comparison grouped by gain or loss of expression in humans. **h**, Average human–chimpanzee differential expression (DE) patterns along the trajectory from pluripotent cells to cortical neurons shown for the seven clusters of differentially expressed genes. NE, neuroepithelium; IPC, intermediate progenitor cell; Norm. FC, normalized fold-expression change. **i**, **j**, Distinct humanspecific pseudotemporal expression patterns (**i**, cluster number with number of genes in parenthesis) and exemplary genes (**j**) for each of the seven clusters.

(33.8%). Differential accessibility peaks are enriched for distinct biological processes (Fig. 3f), and many differential accessibility peaks are accessible specifically in organoids (Fig. 3g) and drive reporter expression in the mouse developing forebrain [277] (Extended Data Fig. 9). Most differential accessibility peaks were located in non-protein-coding regions distal to the promoter (Fig. 3h). The majority of genes that were differentially expressed between human and chimpanzee along the dorsal telencephalon trajectory have one or more differential accessibility peaks nearby (Fig. 3h), and are more likely to be near a differential accessibility region than non-differentially expressed genes (Extended Data Fig. 9). Differential accessibility peaks are also significantly enriched for single nucleotide changes (SNCs) that are fixed in humans and distinct from other primates (Fig. 3i), some of which are predicted to generate new or disrupt existing binding sites for organoid-expressed transcription factors (Extended Data Fig. 9). We annotated differential accessibility peaks and identified organoid-specific potential regulatory regions near differentially expressed genes that have human-fixed SNCs, have undergone accelerated evolution in humans (human accelerated regions (HARs)) or overlap conserved regions that have been deleted in humans (hCONDELs) (Fig. 3j, Supplementary Table 11). This analysis identified many regions specific to NPCs and neurons (Extended Data Fig. 9), and we found 62 HARs that overlap differential accessibility peaks, one of which is near cadherin 7 (*CDH7*), a gene with human-specific expression in neurons (Fig. 3k).

Finally, we wanted to know whether the human-specific gene expression patterns observed in the developing cortex persist into adulthood. We generated snRNA-seq data from post-mortem prefrontal cortex tissues of three humans, two chimpanzees, one bonobo and three macaques by isolating nuclei from sequential sections from basal to apical positions [141] (Fig. 4a). Nuclei clustered into different populations of neurons, astrocytes, oligodendrocytes, microglia and endothelial cells (Fig. 4a, b, Extended Data Fig. 10). Genes specifically expressed in neurons showed higher sequence conservation than genes specific to other cell types (Fig. 4c), astrocytes had the largest number of human-specific differentially expressed genes (Fig. 4d), and some genes with human-specific differential expression in excitatory neurons showed layer specificity (Fig. 4e). A substantial fraction of the genes expressed in excitatory or inhibitory neurons in the adult cortex were also detected in their counterparts in the organoid telencephalon (Fig. 4f, Extended Data Fig. 10). Notably, relatively few genes were commonly detected as differentially expressed (human versus chimp) in organoid and adult excitatory (53 of 354) or inhibitory (13 of 217) neurons (Fig. 4g), and differentially expressed genes detected ubiquitously in the organoid cortex show stronger consistency with adult neurons compared with organoid NPC- or neuron-specific genes (Extended Data Fig. 10). Examples of human-specific differentially expressed genes in adult cell classes (Fig. 4h) include differentially expressed genes detected in developing neurons, such as *COL6A1*, which has been shown to have a protective role in ageing neurons [279]. These results suggest that, with some exceptions, transcriptome differences between human and chimpanzee cortical cells

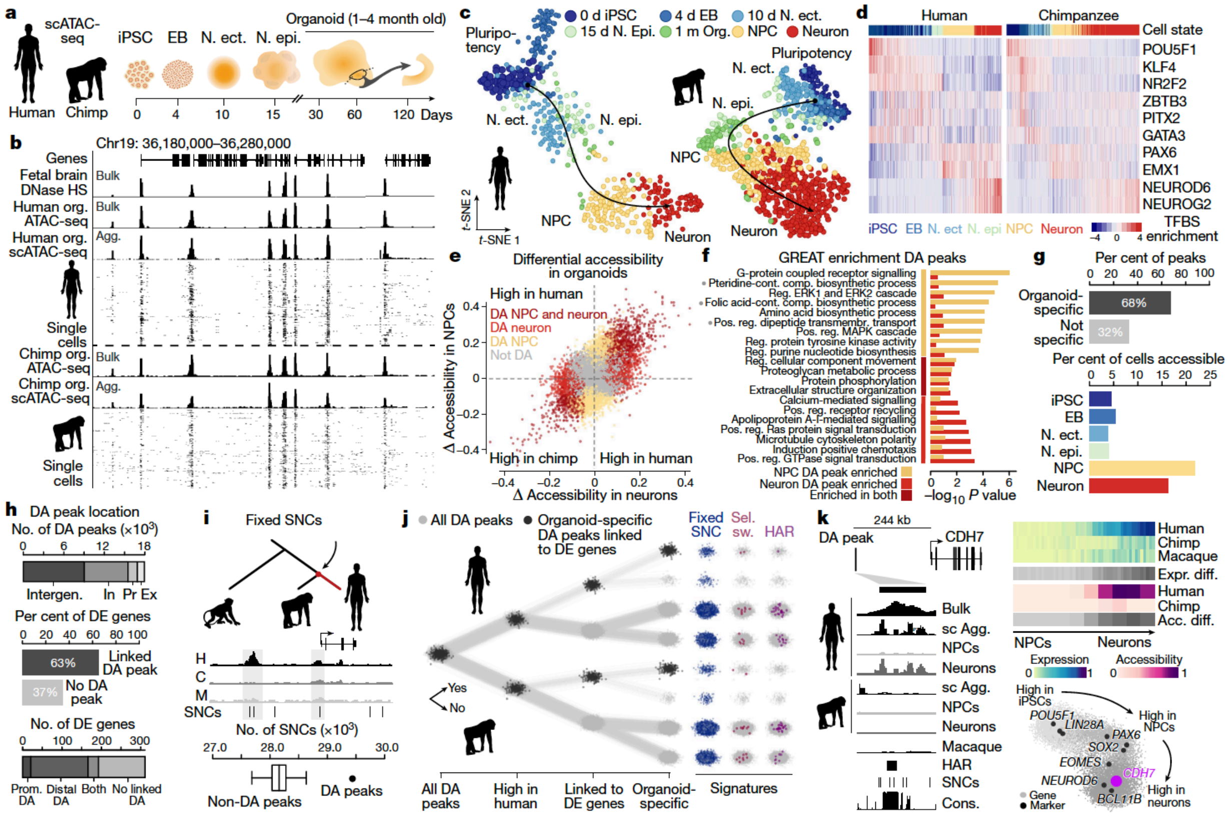


Fig. 3: scATAC-seq reveals chromatin accessibility dynamics during cortex development and evolution. **a**, scATAC-seq was performed at different time points of human and chimpanzee cerebral organoid development from pluripotency to four-month-old organoids (microdissected cortical regions). **b**, Bulk, single-cell and aggregated (Agg.) single-cell ATAC-seq profiles from two-to-four month-old organoids at a representative locus. **c**, ChromVAR [278]-based *t*-stochastic neighbour embedding (*t*-SNE) of human (left, 518 cells) and chimpanzee (right, 908 cells) with cells coloured by time point or cell state. **d**, Heat maps showing binding-motif enrichment for selected transcription factors in cells ordered in pseudotime. **e**, Differentially accessible (DA) peaks detected between human and chimpanzee NPCs and neurons. **f**, Enrichment of selected biological-process gene ontology (GO) terms associated with differentially accessible peaks in NPCs (gold) or neurons (light red) relative to all accessible organoid peaks. Grey dots indicate significantly enriched terms (hypergeometric test, FDR <0.05 and twofold region-based enrichment). Cont., containing; comp., compound; reg., regulation; pos., positive. **g**, Top, percentage of differentially accessible peaks that are accessible only in organoids (organoid specific) or also at earlier stages (not specific). Bottom, percentage of human cells from each cell state accessible at differentially accessible peaks. **h**, A differentially accessible peak was linked to the nearest expressed gene within 1 Mb of that gene's transcription start site. Top, number of differentially accessible peaks located in exonic (Ex), promoter (Pr), intronic (In) or intergenic (Intergen) regions. Middle, percentage of differentially expressed genes linked with differentially accessible peaks. Bottom, proportion of differentially expressed genes with a differentially accessible peak at the promoter region (Prom. DA), distal region (Distal DA), both, or no linked differentially accessible peak. **i**, Numbers of SNCs derived and fixed in humans overlapping differentially accessible peaks and non-differentially accessible peaks (randomly sampled 2,000 times to match the number and average accessibility of differentially accessible peaks). DE, differentially expressed. **j**, Differentially accessible peak annotation noting overlap with selective sweeps (sel. sw.) or HARs. **k**, A differentially accessible peak close to the differentially expressed gene *CDH7* overlaps fixed SNCs and a HAR. Bottom right, a gene correlation network with *CDH7* highlighted. Expr., expression; Acc., accessibility; sc Agg., single-cell aggregated; cons., conservation.

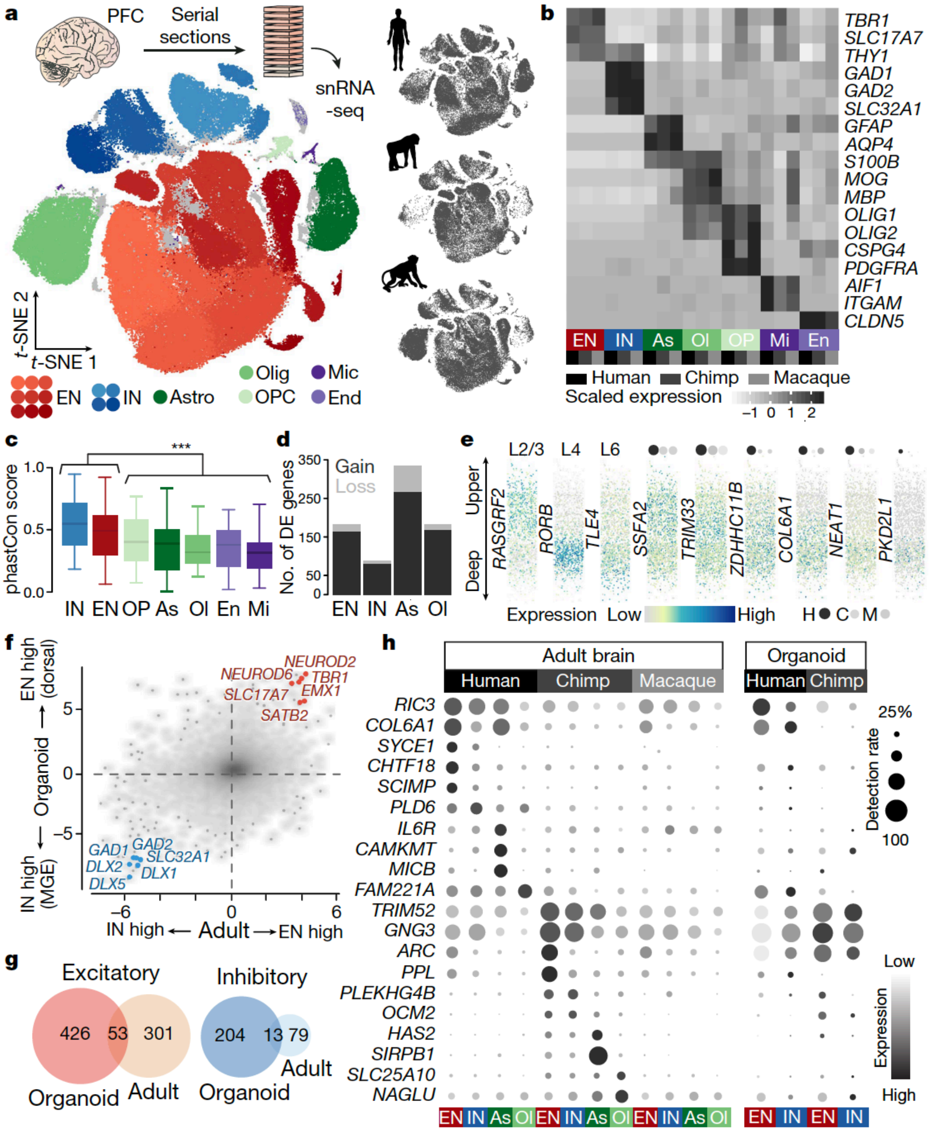


Fig. 4: Adult cortex snRNA-seq reveals shared and distinct patterns of gene expression change compared to organoids. **a**, snRNA-seq was performed on sliced tissue cubes dissected from adult human, chimpanzee or bonobo, and macaque frozen prefrontal cortex tissue. Projection of integrated data shows different clusters of major cell classes in the different species (50,035 human, 33,847 chimpanzee or bonobo, and 50,403 macaque nuclei). Astro, astrocyte; Olig, oligodendrocyte; OPC, oligodendrocyte precursor cell; Mic, microglia; End, endothelial cell. **b**, Average cluster expression separated by species reveals similar patterns of marker gene expression for seven cell classes. As, astrocyte; Ol, oligodendrocyte; OP, oligodendrocyte precursor; Mi, microglia; En, endothelial cell. **c**, Genomic conservation (average phastCon scores) of markers for seven cell classes (***$P < 0.0001$, two-sided Wilcoxon's rank-sum test). Boxes represent IQR and whiskers represent 1.5 x IQR. **d**, Number of human-specific differentially expressed genes in each of the four major cell classes. **e**, Predicted laminar expression patterns of three layer (L) markers (left) and six human-specific differentially expressed genes (right) in adult excitatory neurons with size and shade of dots above the expression pattern plot showing detection rates and average expression levels, respectively, in the three species. **f**, $\log_2$(fold change) between excitatory and inhibitory neurons shows consistency between organoid dorsal and ventral forebrain and adult tissue. **g**, Overlap of human–chimpanzee differentially expressed genes detected in adult neurons and organoid trajectories. **h**, Dot plot showing expression patterns and detection rates across adult and organoid cell classes for human-specific differentially expressed genes in adult cell classes.

are linked to developmental stages.

In sum, our data illuminate expression and regulatory features of cell states that are uniquely human and provide an extensive resource to guide exploration into the gene regulatory mechanisms that distinguish the developing human and chimpanzee brains, some of which persist into adulthood.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1654-9.

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Pluripotent stem cell lines and organoid culture. We acquired 6 human induced pluripotent stem cell (iPSC) lines (Sojd3, Hoik1, Kucg2, Wibj2 from the HipSci resource [280]; h409b2 from the RIKEN BRC cell bank [114]; Sc102a1 from System Biosciences), one human ES cell line (H9, WiCell) [281], three chimpanzee iPSC lines (SandraA15; PR818-515, originally generated by the Gage laboratory and kindly provided to us by the R. Livesey group; JoC, generated in this study), one bonobo iPSC line (Bokela, generated in this study), one ES macaque cell line (MN1 [59], kindly provided through the R. Livesey group from E. Curnow) and one orangutan cell line (Toba [222]). The iPSC line JoC (chimpanzee, Tchimpounga Sanctuary) was reprogrammed from blood cells (primary lymphocytes) using plasmid based reprogramming [81] and Bokela (bonobo, Zoo Leipzig) was reprogrammed from fibroblasts using the StemMACS mRNA transfection kit (Miltenyi Biotec). Cell lines were validated for pluripotency markers by immunhohistochemical stainings using the Human Pluripotent Stem Cell 3-Colour Immunohistochemistry Kit (R&D Systems, SC021) and were differentiated into the three different germ layers using the Human Pluripotent Stem Cell Functional Identification kit (R&D Systems) and StemMACS Trilineage Differentiation Kit (Miltenyi Biotec). Karyotyping was carried out using Giemsa banding at the Stem Cell Engineering facility, a core facility of CMCB at Technische Universität Dresden, and karyotypes were found to be normal. Cell lines were cultivated using standard feeder-free conditions in mTeSR1 (StemCell Technologies) or StemMACS iPS-Brew XF (Myltenyi Biotec) on matrigel-coated plates and differentiated into cerebral organoids using a whole organoid differentiation protocol [250] (Supplementary Table

1). Toba orangutan iPS cells were maintained on feeder cells and cultivated in ReproMix (1:1 mix Primate ES medium (Reprocell) and TeSR2 (StemCell Technologies), 8 µg/ml bFGF (Sigma), 50 U/ml Penicillin/Streptomycin (ThermoFisher)). Colonies were picked from feeder cells to collect cells for generating EBs. Cell lines were regularly tested for mycoplasma using PCR validation (Venor GeM Classic, Minerva Biolabs) and found to be negative. Permission to work with human and non-human primate iPSC lines was obtained through the Sächsisches Staatsministerium fur Umwelt und Landwirtschaft (Az.: 55- 8811.72/26, Az.: 55-8811.72/26/382, Az.: 55-8811.72/26/393, 54-8452/26/7). The utilization of hESCs for the generation of cerebral organoids was approved by the ethics committee of the Robert Koch Institut (https://www.rki.de/DE/Content/Gesund/Stammzellen/Register/reg-20161027-Paeaebo.html) as well as by the Ethics committee of northwest and central Switzerland (2019-01016) and the Swiss federal office of public health.

Single-cell RNA-seq data generation. A summary of all single-cell experiments can be found in Supplementary Table 1. For organoid experiments (1 month, 2 months, 3 months, 4 months), whole organoids were dissociated for generating single cell gene expression libraries. In brief, organoids were transferred to HBSS (without Ca^{2+} and Mg^{2+},-/-) and cut into two pieces to clear away debris from the centre of the organoid (2–3 washes in total). Organoid pieces were then dissociated using Neural dissociation kit (P) using Papain-based dissociation (Miltenyi Biotec). Organoid pieces were incubated in Papain at 37 °C (enzyme mix 1) for an initial 15 min, followed by addition of Enzyme A (enzyme mix 2) to the Papain mix. Organoid pieces were then triturated using wide bore 1,000-ml tips and incubated for additional intervals of 5-10 min with triturations between the incubation steps, amounting to a total Papain incubation time of approximately 45 min. Cells were filtered through a 30 µm strainer and washed, centrifuged for 5 min at 300 g and washed 3 times with HBSS (-/-). Cells were then analysed using Trypan Blue assay, counted using the automated cell counter Countess (Thermo Fisher), and diluted for an appropriate concentration to obtain approximately 6,000 cells per lane of a 10x microfluidic chip device. Typically, cells from one organoid were loaded per lane in the microfluidic device, and in some cases organoids from different lines were pooled onto the same lane and demultiplexed based on single-nucleotide polymorphisms. For 1 month organoids, three pooled 409b2 and one H9 organoid were dissociated and cells from the two cell lines were mixed at equal ratios to be loaded on the chip. For a set of 2 month HipSci organoid data, organoids were dissociated for all four HipSci cell lines and pooled at equal ratios to be loaded on one lane of the microfluidic device aiming for 10k cells. Fluidigm C1 data (Supplementary Table 1) were generated as previously described [222] and cells from chimpanzee SandraA 75d organoids were microdissected regions from vibratome slices for which single cell suspensions were generated as described above. Single cells were then sorted into 96-well plates using a FACS Aria III sorter and further processed using the SmartSeq2 protocol [195] to generate cDNA and the NexteraXT kit (Illumina) to generate sequencing libraries. All libraries (10x and Fluidigm C1/SmartSeq2) were sequenced on Illumina's Hiseq2500 platform in paired-end mode (100 bp Fluidigm C1/SmartSeq2; 26+8 bp, 100 bp 10x).

Early stages of organoid differentiation (iPS cells to neuroepithelium). For iPSC/ESC single-cell experiments, cells were detached from cell culture dishes using TrypLExpress (Thermo Fisher) incubation for 5 min followed by addition of mTeSR1. Cells were centrifuged for 5 min at 200g and resuspended in mTeSR1, filtered through a 20 µm strainer and washed with mTeSR1. Cells were then centrifuged again for 5 min at 200g and resuspended in mTeSR1, counted, diluted to the same concentration and mixed at equal ratios for the three cell lines to be loaded on the 10x microfluidic chip aiming for 10,000 cells. Thirty embryoid bodies (EBs), 7–15 neuroectoderms, and 1–3 neuroepithelium of each cell line were pooled for each dissociation, respectively. Cells were obtained by papain dissociation as described above for organoid dissociation, with slightly shorter incubation times in enzyme mix 1 (approximately 30 min.). For 10x experiments, cells from the three different cell lines were diluted and mixed at equal ratios to be loaded on the microfluidic chip device.

Single-cell experiments were conducted using the 10x Chromium Single Cell 3' v2 Kit following the manufacturer's instructions. In brief, cells were mixed with reverse transcription mix, gel beads and oil were loaded on the chip device to be coencapsulated into droplets, which underwent first strand cDNA synthesis thereby tagging mRNAs with a UMI and a unique cell barcode. All following steps were conducted in bulk by breaking the droplets and cleaning up and amplifying the cDNA. Single-cell libraries were then constructed by fragmentation, end repair and adaptor ligation and amplification using library specific index sequences as provided by 10x Genomics. Quantification and quality control of libraries was performed using High Sensitivity DNA assays for Agilent's Bioanalyzer and sequenced on a HiSeq2500 in Rapid or HighOutput sequencing mode. Typically, one 10x library was sequenced on one lane of a sequencing flow cell, with the exception of the HipSci organoids for which three pooled libraries (each library contained pooled cells from four dissociated HipSci organoids from different cell lines) were sequenced on two lanes of a flow cell. See Supplementary Table 1 for more details.

Immunohistochemistry Organoids were washed in PBS before fixing in 4% PFA for 2–4 h. The excess of fixative was removed with three PBS washes and organoids were then transferred to a 30% sucrose solution for 24–48 h for cryoprotection. Finally, organoids were transferred to plastic cryomolds (Tissue Tek) and embedded in OCT compound 4583 (Tissue Tek) for snap-freezing on dry ice. For immunohistochemical stainings, organoids were sectioned in slices of 20 µm thickness using a Leica CM3050 S cryostat and Microm HM 560 (Thermo Fisher Scientific) at -15 to -20 °C. Organoid sections were quickly washed in PBS to remove any residual OCT. Then, sections were incubated in antigen retrieval solution (HistoVT One, Nacalai Tesque) at 70 °C for 20 min. Excess solution was washed away with PBS and the tissue was incubated in blocking-permeabilizing solution (0.3% Triton, 0.2% Tween-20 and 5% normal goat serum in PBS) for 1h at room temperature. Afterwards, sections were incubated overnight at 4°C in blocking-permeabilizing solution containing antibodies anti-PAX6 (mouse, 1:1,000, Thermo Fisher Scientific, MA1-109; rabbit, 1:300, Covance, PRB-278P)and anti-CTIP2 (rat, 1:1,000, Abcam, AB18465), anti-SATB2 (rabbit, 1:500, Abcam, Ab92446; mouse, 1:500, Abcam, Ab51502), anti-Tbr2 (mouse, 1:500, MPI-CBG Antibody Facility [61]). On the next day, sections were rinsed three times in PBS before incubation for 1 h at room temperature in secondary antibody solution, which contained blocking-permeabilizing solution, DAPI (1:3000), Alexa Fluor 488-conjugated anti-rabbit antibody (goat, 1:1,000, Thermo Fisher, A11008), Alexa Fluor 546-conjugated anti-mouse antibody (goat, 1:500, Thermo Fisher Scientific, A-21123), Alexa Fluor 647-conjugated anti-rat antibody (goat, 1:500, Thermo Fisher Scientific, A-21247) and Alexa Fluor 488-conjugated anti-mouse (A21202) and anti-rat antibody (A21208), Alexa Fluor 555-conjugated anti-rabbit antibody (A31572), Alexa Flour 647-conjugated anti-mouse antibody (A31571) (all donkey-derived, 1:500, Molecular Probes). Finally, remainders of secondary antibody solution were washed off with PBS before covering with ProLong Gold Antifade Mountant medium (Thermo Fisher Scientific). Stained organoid cryosections were imaged using a confocal laser scanning Olympus Fluoview FV1200 microscope and Zeiss LSM 880 Airy upright microscope. Whole-section tilescans composed of 3 different z-plane images (z-step = 5-8 µm) were acquired using a 10x magnification objective, Plan-Apochromat 10x/0.45 M27 and Plan-Apochromat 20x/0.8 M27 objectives. Images were then stitched, stacked and further processed using the Olympus Fluoview 4.2b software and ImageJ (Fiji).

Single-cell RNA-seq data preprocessing. We used Cell Ranger, the set of analysis pipelines suggested by 10x Genomics, to demultiplex raw base call files to FASTQ files and align reads to the human genome and transcriptome (hg38, provided by 10x Genomics) with the default alignment parameters. Pooled samples, including samples from different species or human lines, were then demultiplexed using a two-step procedure based on the read mapping results. In the first step, the genome alignment between human (hg38) and chimpanzee (panTro5) was downloaded from UCSC Genome Browser. Sites with diverged bases between human and chimpanzee were obtained based on the genome alignment. Reads covering the species-diverged sites were collected for each reported cell, with the number of bases matching each species counted. Cells with more than 80% reads covering the species-diverged sites matching with one species were assigned as cells from the species. For those samples with human cells from different lines pooled, a second step of demultiplexing was done using demuxlet [282], based on the genotyping information of lines downloaded from HipSci websites (Kucg2, Wibj2, Hoik1, Sojd3) or called using bcftools based on the unpooled scRNA-seq data (H9, 409b2). Cells with the best singlet likelihood no less than 50 higher than the second best singlet likelihood and estimated mixture ratio less than 30% were labelled as their bestmatched lines. All cells failing to pass any of the above threshold were classified as doublets and excluded from the following analysis.

Seurat [283] was then applied for further data processing. Cells with more than 6,000 or less than 200 detected genes, as well as those with mitochondrial transcripts proportion higher than 5% were excluded. After the log-normalization, confounding factors including the number of detected genes and proportions of mitochondrial transcripts were also regressed out. Highly variable genes were then obtained as genes with dispersion higher than 0.5 and normalized expression level between 0.0125 and 3, followed by principal component analysis (PCA) based on the z-transformed expression levels of the identified highly variable genes (Supplementary Table 2). The top-20 principal components (PCs) were used to do clustering (with a resolution of 0.6) using Seurat. Additional quality controls of the measured cells were based on primary cell type predictions by using public human fetal brain scRNA-seq data (Nowakowski dataset)[219]. In brief, a Lasso logistic regression model was built, using gene expression ranks of the Nowakowski dataset as the training set, to predict the primary cell type identity of each single cell in two-month-old and four-month old organoids. Cells which were predicted to be of 'glycolysis' identity were excluded, so as cells in the Seurat clusters where more than 80% of cells were predicted as of 'glycolysis' identity. Heterogeneity analysis of human (Extended Data Fig. 2, Supplementary Tables 3, 4) and chimpanzee (Extended Data Fig. 4, Supplementary Tables 5, 6) full lineage data was performed using *t*-SNE based on the top PCs identified (top 20 PCs for human, top 15 PCs for chimpanzee). Cluster identities were assigned based on cluster gene markers (Supplementary Tables 4, 6) as determined by FindAllMarkers function in Seurat (min percentage of cells expressed = 0.25 and log fold change threshold = 0.25) and gene expression of known marker genes. For human data, cells from 409b2 and H9 and were integrated using canonical correlation analysis (CCA) as implemented in Seurat (v.3). In brief, data were normalized and the top 2,000 highly variable genes for 409b2 and H9 cells were determined using the vst method. The datasets were integrated based on the top 20 CCs using the Seurat method by identifying anchors and integrating the datasets. The resulting integrated data were scaled and PCA was performed. Clustering was performed based on the top 20 PCs and using a resolution of 0.6. Feature plots show non-integrated expression values. Cluster markers were determined using Wilcoxon test considering only genes that show a minimum log fold expression change of 0.25 in at least a fraction of 0.25 of cells in the clusters using the non-integrated expression values.

RSS and construction of pseudocell transcriptomes. RSS of one cell to the Human Developing Brain atlas (BrainSpan) was defined as the normalized similarity between gene expression levels of the cell and gene expression levels of each of the 237 fetal samples with RNA-seq data in the BrainSpan database in Allen Brain Atlas. To increase discrimination of different reference samples, only the highly variable genes of the BrainSpan dataset (see Supplementary Table 2), defined based on expression variation-mean comparison of the reference dataset, were used for the RSS calculation. Between each cell and each sample in the

BrainSpan dataset, Pearson correlation coefficient (PCC) was calculated across the BrainSpan-highly variable genes. z-Transformation was then applied to PCCs between each cell to the 237 fetal BrainSpan samples to get the normalized similarities. To construct pseudocells, single cells were first grouped based on their sample sources and Seurat clusters. Within each group of cells, that is, those cells from the same sample and in the same Seurat cluster, cells were selected randomly with a selection probability of 20%. The selected cells were called pseudocell seeds or territory capitals. The ten nearest neighbours of each seed, based on Euclidean distances of the top-20 PCs, were then assigned to the seed, forming a pseudocell territory. If one cell was assigned to multiple pseudocell territories, one territory was chosen randomly. The expression level of one gene in each pseudocell was then calculated as the average gene expression level across cells in the pseudocell territory.

Visualization, lineage identification and pseudotime estimation of pseudocells for reconstructing human cerebral organoid differentiation from pluripotency. First, PCA was applied to a pseudocell expression matrix using the z-transformed expression levels of the highly variable genes as input. Euclidean distance between the top 10 PCs of each pair of pseudocells was calculated and a k-nearest neighbour (kNN) network ($k = 100$) was then calculated with the constraint to only consider pseudocells from the same or nearby stages when screening for nearest neighbours. The kNN network was visualized using SPRING [274]. To construct the pseudotime course of human cerebral organoid differentation from pluripotency, the Walktrap community identification algorithm (implemented in the R package igraph) was applied to the above kNN network to identify network communities. The resulting communities were manually aggregated into four groups to minimize branches in each group. A diffusion map algorithm (implemented in R package destiny [225]) was applied to pseudocells in each of the four groups, with the expression levels of the highly variable genes of pseudocells as the input. The ranks in DC1 were used as the pseudotimes. We used an F-test based ANOVA analysis to identify genes with pseudotime-dependent expression patterns. In brief, we established a natural splined linear regression model (ns function in the R package splines) with six degrees of freedom (df), with expression levels as the response variable and pseudotimes as the independent variable, for each of the highly variable genes. An F-test was applied, to compare variation explained by the splined linear model with that of the residuals normalized by degrees of freedom. Bonferroni correction was performed across tested genes, with a corrected P value threshold 0.01 to identify genes with pseudotime-dependent expression. The analysis was applied to the four groups of pseudocells separately.

Visualization, lineage identification, pseudotime estimation of cells in human two–month-old cerebral organoids from different individuals. Pseudocells were constructed for the human two-month-old organoids as above with constraint on samples and based on cells with predicted primary cell types as one of radial glia, intermediate progenitor, excitatory neuron, and inhibitory neuron. RSS to the BrainSpan fetal samples was calculated for each pseudocell, with distance between two pseudocells defined as the correlation distance between RSS of the two pseudocells. The kNN network ($k = 20$) was then constructed and SPRING was used to determine coordinates of pseudocells for visualization. To further discriminate pseudocells representing different neuronal lineages, a Walktrap algorithm for network community identification was applied to the RSS-based kNN ($k = 100$). Communities that were significantly connected and showing concordant marker expression or similarity spectrum were aggregated, which resulted in three progenitor-to-neuron trajectories. Based on gene expression level ranks of cells in the three defined trajectories, two Lasso logistic regression models were trained, one for classification of cortical and ventral lineage, while the other one for classification of all the three trajectories. The first model was applied to pseudocells in the community C6 which was significantly connected to cortical and ventral trajectories, while the second model was applied to pseudocells in the community C4, which was significantly connected with both the non-telencephalon pseudocells and community C4. With a unique lineage label defined for each pseudocell, a 1 x 500 self-organizing-map (SOM) model was trained for each of the three trajectories, using RSS of pseudocells within the lineage as the training data. The index of neuron that one pseudocell was assigned to was used as its pseudotime. Diffusion map analysis was also applied to pseudocells at each trajectory, with highly variable gene expression as the input, with ranks of DC1 defined as alternative pseudotime of pseudocells. Pseudotimes obtained by the two methods are highly correlated (Spearman correlation is 0.91 and 0.92 for the dorsal and ventral telencephalon trajectories, respectively).

To project the single-cell data to the cell embedding space that was defined for pseudocells, two support vector regression (SVR) models (implemented in the R package e1071), each of which was for one dimension of the embedding, were trained using RSS of pseudocells as the training set. The trained models were applied to RSS of single cells for their predicted coordinates. Such coordinates were further refined by pushing each cell to its nearest pseudocell with smallest correlation distance of RSS to be 70% closer. Similarly, a support vector machine (SVM) model was trained (implemented in the R package e1071) using RSS of pseudocells for the three trajectories, and applied to RSS of single cells for their trajectory identity. After that, the corresponding SOM model for pseudotime estimation was applied to RSS of each single cell for its estimated pseudotime.

Dynamic time warping-based alignment of pseudotime courses. We used a dynamic time warping (DTW) algorithm to align different pseudotime courses. In brief, each pseudotime course was evenly broke into 50 blocks. Average gene expression levels of pseudocells or cells within each block was calculated. Pairwise distances between blocks from the two courses were calculated as the Pearson correlation distance, that is, 1 - PCC, across the highly variable genes in cells of both pseudotime courses. Suppose $d_{i,j}$ represents the distance between the i-th block in the reference pseudotime course and the j-th block in the query pseudotime course.

We defined **D** as the alignment distance matrix, where

$$\mathbf{D}_{i,j} = min(\mathbf{D}_{i\text{-}1,j}, \mathbf{D}_{i,j\text{-}1}, \mathbf{D}_{i\text{-}1.j\text{-}1}) + d_{i,j}$$

A trace-back procedure was then performed to get the alignment. Three modes of alignment were implemented. In the first mode, the fixed-end alignment, the initialization of **D** was done as:

$$\mathbf{D}_{i,j} = \begin{cases} d_{i,1} & \text{if} \quad j = 1 \\ \mathbf{D}_{1,j\text{-}1} + d_{1,j} & \text{if} \quad i = 1 \end{cases}$$

In the other two modes, the fixed-start and end-to-end alignments, D was initialized as:

$$\mathbf{D}_{i,j} = \begin{cases} d_{1,1} & \text{if} \quad i = 1 \quad \text{and} \quad j = 1 \\ \mathbf{D}_{1,j\text{-}1} + d_{1,j} & \text{if} \quad i = 1 \\ \mathbf{D}_{i\text{-}1,1} + d_{i,1} & \text{if} \quad j = 1 \end{cases}$$

In the trace-back step, a fixed-end and end-to-end alignment was started from $\mathbf{D}_{M,N}$, where M and N are the numbers of blocks at the reference and query pseudotime courses, respectively. On the other hand, the trace-back step was started from $\mathbf{D}_{m,N}$, where $m = \text{argmin}_i(\mathbf{D}_{i,N})$. In our study, the fixed-end alignment was used to align the cortical and ventral lineage pseudotime course of human organoid cells; the fixed-start alignment was used to align pseudotime courses of human and chimpanzee cortical pseudocells; the end-to-end was used in the truncated alignment of pseudotime courses of different species.

Reconstruction of chimpanzee cerebral organoid differentiation from pluripotency. We applied a similar procedure as mentioned above describing the reconstruction of human cerebral organoid differentiation from pluripotency to reconstruct the organoid differentiation trajectory from chimpanzee single-cell RNA-seq data. In brief, the single-cell RNA-seq reads were mapped to the human-chimpanzee-macaque consensus genome and counted using Cell Ranger. Seurat was used for further preprocessing including gene expression normalization, confounding factor regression, PCA and clustering, Cells from organoid samples with predicted primary cell type identity of 'glycolysis', as well as cells within clusters with more than 80% cells having 'glycolysis' identity, were excluded. Pseudocells were then constructed with a seed selection probability of 20% and constraints on samples and Seurat clusters. PCA was applied to expression levels of highly variable genes across pseudocells, and pairwise distances of pseudocells were calculated as the Euclidean distances between the top-10 PCs. The kNN network ($k = 100$) of pseudocells was constructed, linking every pseudocell with its 100-nearest pseudocells representing the same or nearby stages. Three-month-old and four-month-old organoids were seen as the same stage. The Walktrap network community identification algorithm was applied and the resulted community labels (walktrap communities) of pseudocells were compared with the predicted community labels (projected communities) based on a Lasso logistic regression model trained by ranks of gene expression levels of the human pseudocells representing the human organoid differentiation from pluripotency as described above. Any walktrap community with $<$1,000 kNN connections with other communities was discarded. One of the four labels: early, cortical, ventral, non-telencephalon was assigned to one walktrap community if more than 95% of pseudocells within the community were from the same group according to their projected communities. For one community with more than 10% of pseudocells with projected communities belonging to both ventral and midbrain-hindbrain groups, the non-telencephalon label was only assigned to pseudocells with projected communities in the non-telencephalon group. The diffusion map algorithm was applied to each of the four pseudocell groups, using the expression levels of highly variable genes as input, to estimate their pseudotimes. For the cortical, ventral and midbrain-hindbrain groups, the ranks of DC1 was used as the pseudotimes. For the early group, a principal curve (implemented in the R package princurve) was fitted in the DC1–DC2 space. The order of pseudocells projecting to the resulted principal curve was used as the pseudotimes.

Human-chimpanzee-macaque consensus genome. The construction of the consensus genome was performed using the procedure as described [141, 156]. In brief, the chained and netted pairwise genome alignments of the human (hg38) and chimpanzee (panTro5) genomes, and the human and macaque (rheMac8) genomes, were downloaded from UCSC Genome Browser. Based on the downloaded pairwise genome alignments, a multiple genome alignment of human-chimpanzee-macaque was constructed using multiz. Based on the human-chimpanzee-macaque genome alignment, we constructed the consensus genome by masking all discordant sites including mismatches, insertion/deletion (indels), as well as 6-bp flanking regions of indels on the human genome. The obtained consensus genome was indexed with gene annotation in GENCODE v.27 for read mapping to the consensus genome with Cell Ranger.

Pseudotime estimation of cerebral organoid cells in different species. Single cell RNA-seq data of organoids with ages from two-month-old to four-month-old in human, chimpanzee, and macaque were mapped to the human-chimpanzee-macaque consensus genome and counted using Cell Ranger. Further preprocessing using Seurat was applied separately for data from the three species. Only cells with predicted primary cell type identities as radial glia, intermediate progenitors, excitatory neurons, or inhibitory neurons were included in the later analysis. Pseudocells were constructed for humans and chimpanzees, both with a coarse grain ratio of 20% and constraints on samples and Seurat clusters. The RSS to the BrainSpan dataset was calculated for each pseudocell, and the SVM model for lineage estimation was applied to estimate the lineage identity of each pseudocell. Focusing on the cortical lineage, a diffusion map analysis was applied to cortical pseudocells of the three species, respectively. The ranks of DC1 were used as the pseudotimes of the pseudocells. In macaque, similar procedure was applied directly to single cells without pseudocell construction.

Truncated DTW-based alignment of pseudotime courses representing neural progenitors and deeper layer neurons in different species. We used the first DC discriminating *BCL11B*$^+$ and *SATB2*$^+$ cortical neurons (DC3 in chimpanzee, DC4 in macaque) to identify upper layer (UL) neurons, as the pseudocells in the branch with highest expression level of *SATB2*. To identify potential upper layer neurons in human, we first retrieved markers of upper and deeper layer (DL) excitatory neurons [219]. The sum expression levels of UL and DL markers was then calculated for each pseudocell in human and chimp, with the UL-specificity score (s_{UL}) being defined as the UL/DL markers expression ratio. The distribution of (s_{UL}) in UL neurons in chimpanzee was used to determine the threshold to discriminate UL neurons from other cell types ($s_{\mathrm{UL}} > 0$). All UL neurons in the three species were excluded from the following analysis.

To correct for the DL neuron maturation timing differences between human and the other two species, a two-step pseudotime course alignment strategy was used. The first step, namely the trimming step, aims to determine the pseudotime points in chimpanzee and macaque which correspond to the latest pseudotime point in human. In brief, an SVR model with Gaussian-kernel was first constructed, with chimpanzee or macaque pseudotimes as the response variables and the RSS as the dependent variables. Two models were trained with the chimpanzee pseudocells and macaque cells respectively, and applied to the human pseudocells to predict their corresponding chimpanzee and macaque pseudotime points. Two constrained B-splines regression models (F_{HC}, F_{HM}) were then fitted (implemented in the R package cobs): human pseudotimes of human pseudocells ($\widehat{t_{\mathrm{h}}}$) versus their predicted chimpanzee ($\widehat{t_{\mathrm{c}}}$) or macaque (t_{m}) pseudotimes, with constraints of $F_{\mathrm{HC}}(t_{\mathrm{h}} = 0) = F_{\mathrm{HM}}(t_{\mathrm{h}} = 0) = 0$. $F_{\mathrm{HC}}(t_{\mathrm{h}} = 1)$ and $F_{\mathrm{HM}}(t_{\mathrm{h}} = 1)$ were used as the pseudotime thresholds to select chimpanzee pseudocells and macaque cells. Chimpanzee DL neurons with pseudotime $t_{\mathrm{C}} > F_{\mathrm{HC}}(t_{\mathrm{h}} = 1)$, as well as macaque DL neurons with pseudotime $t_{\mathrm{M}} > F_{\mathrm{HM}}(t_{\mathrm{h}} = 1)$, were excluded in the following analysis. The second step, namely the alignment step, was then applied to the remaining pseudotime courses of the three species. An end-to-end DTW-based alignment, as described above, was used to align the human pseudotimes with pseudotimes of each of the other two species using the human pseudotime course as the template.

Identification, clustering, and species specificity of differentially expressed genes between humans and chimpanzees. Human genes resulting from duplication or rearrangement that do not exist in other apes were collected via literatures [61, 63, 65, 284]. In total, 41 genes were obtained, 24 of which were included in the human gene annotation in Cell Ranger. 23 of them were detected in at least one human cell, with *FCGR1B* being the exception. Expression patterns of those genes were examined and summarized manually (Supplementary Table 7).

To compare transcriptome changes of the developmental trajectory from cortical neural progenitors to deeper layer neurons between human and chimpanzee, an F-test based comparison was applied to the expression profile along pseudotimes of the two species. In brief, for each gene, a natural spline linear regression model (df = 6) was constructed for human and chimpanzee pseudocells along the aligned pseudotime course, without discriminating human and chimpanzee samples, and used as the null model (m_0). The alternative natural spline linear regression model was also constructed, with each species having its own slopes and intercept (m_1). The residuals of the variation, which cannot be explained by each model, were compared by an F-test. Non-ribosomal genes with BH-corrected $P < 0.01$ were identified as differentially expressed (DE) genes between human and chimpanzee along the developmental trajectory from cortical neural progenitors to deeper layer neurons (Supplementary Table 8).

To estimate the robustness of the identified differential expression to the number of used lines, as well as the pseudocell distribution along the pseudotime course, we used a series of replaceable pseudocell sampling procedure with constraints. In brief, in each round of replaceable pseudocell sampling, the candidate pseudocells to be selected are restricted to be those from a certain number of human cell lines. In addition, the subsampling in human pseudocells is performed to recapitulate the pseudocell distribution along the aligned pseudotime of chimpanzee pseudocells, that is, each of the ten pseudotime bins contains the same number of human and chimpanzee pseudocells. This sampling procedure was performed 100 times for each possible number of human lines, ranging from one to seven. Differential expression analysis, as described above, was applied to compare gene expressions of human pseudocells in each sampling with the chimpanzee pseudocells. Robust DE genes were determined as DE genes which can be detected in at least 80% of tests performed with replaceable pseudocell samplings with any number of used human cell lines.

A similar strategy was also used to estimate the false positive human-chimpanzee DE genes due to differences between cell lines. In each sampling, two lines were randomly selected as group one, and a certain number of lines, ranging from one to five, were selected from the remaining lines as group two. For each group, pseudocells were randomly sampled from the selected lines to recapitulate the pseudocell distribution along the aligned pseudotime of chimpanzee pseudocells. Such sampling was performed 100 times for each possible number of lines used in group two. The transcriptome trajectory from cortical neural progenitors to deeper layer neurons in macaque organoids was used as the evolutionary outgroup to determine species specificity of the identified human-chimpanzee DE genes. First, the cumulative expression divergences of each gene between human and macaque (d_{HM}), and between chimpanzee and macaque (d_{CM}), were calculated. The cumulative expression divergence was calculated by summing up absolute values of average expression differences between species at the 50 pseudotime bins of equal sizes along the aligned pseudotimes. The human-chimpanzee DE of one gene is seen as human-specific if $d_{\mathrm{HM}} - d_{\mathrm{CM}} > \max(d_{\mathrm{HM}}, d_{\mathrm{CM}})/2$. Genes with chimpanzee-specific differential expression were identified in the same way. Genes with human-specific differential expression were then clustered based on their human–chimpanzee DE along pseudotimes. Average expression differences between human and chimpanzee at each of the 50 pseudotime bins along the pseudotimes was calculated for each gene with human-specific differential expression (denoted as d_{t} at pseudotime bin t), and then normalized as $\widehat{d_{\mathrm{t}}} = d_{\mathrm{t}}/\left(\underset{t}{\max} d_t - \underset{t}{\min} d_t\right)$. Hierarchical clustering (Ward algorithm) was then used to cluster those genes into nine clusters, with distances between genes calculated as the Euclidean distances between their normalized differential expression spectrums. Clusters with fewer than five genes were discarded. We annotated genes with human-specific expression patterns using the *Homo sapiens* Gene Ontology Annotation file (validation date: 21 April 2017) provided by the Gene Ontology Consortium.

Processing of the Fluidigm C1 based scRNA-seq data of cerebral organoids. In addition to the newly generated Fludigm C1 (SmartSeq2)-based scRNA-seq data, we further retrieved published sequencing data of 786 and 344 single cells from human and chimpanzee cerebral organoids [114, 222], in the format of FASTQ files from GEO accession numbers GSE75140 and GSE86207 (CMK dataset). All the reads were mapped to the human-chimpanzee-macaque consensus genome using STAR (v.2.6.1d) with '–quantMode' parameter set to TranscriptomeSAM and GENCODE v.27 annotation provided. Gene expression levels in each cell were quantified as TPM by RSEM (v.1.3.1). Additionally, we retrieved the recently published gene expression matrix representing 3,211 cells from human and chimpanzee cerebral organoids (excluding redundant cells from GSE75140 and GSE86207) and 4,854 cells from human and macaque fetal brains [20].

Based on the resulting gene expression profile, RSS to the fetal Brainspan dataset was calculated as described above for each cell, with 248 genes with significant differential expression between cortical neurons measured by Smart-seq and Smart-seq2 excluded from the references. Distances between organoid cells were calculated as the Pearson's coefficient distances between RSS of cells. Distances between cells from fetal brains were calculated in the same way. The resulted distance matrices of all organoid cells and fetal brain cells were used as the input to generate t-SNE embeddings. A kNN network ($k = 50$) was generated for organoid cells and fetal brain cells separated based on the RSS-based distances, and a Walktrap algorithm for network community identification was applied to identify cell clusters, which were further annotated based on their marker genes. Based on the cell type annotation, the diffusion map analysis, with the RSS profiles as input, was applied to the dorsal forebrain NPCs and neurons in organoids and fetal brains, respectively. The ranks of DC1 were used as the pseudotimes.

To validate the human-chimpanzee differential expression identified in our droplet-based scRNA-seq data using the C1-based cerebral organoid data, the organoid dorsal telencephalon pseudotemporal trajectory was first split into ten intervals. In each pseudotemporal interval, the human-chimpanzee DE was calculated as the $\log_2$-transformed fold change ($\log_2$FC) between the average expression of human and chimpanzee cells in the interval. Here, the CMK dataset and other datasets which used a distinct quantification method were processed separately. A similar strategy was also applied to the aligned droplet-based human and chimpanzee pseudotemporal trajectories. Generalized $\log_2$-transformed fold change ($g\log_2$FC), defined as the average $\log_2$FC across the pseudotemporal intervals, as well as the maximum $\log_2$FC across the intervals ($m\log_2$FC), was further calculated for each human-chimpanzee robust DE genes in organoids. A DE gene is seen as being consistent in the two datasets if both $g\log_2$FC and $m\log_2$FC of the C1-based and droplet-based human-chimpanzee comparisons are of the same signs (refer as consistent DE genes). The pseudotemporal intervals with the maximum fold change in the droplet-based and C1-based trajectories were also obtained and compared for the consistent DE genes. This procedure was also applied to compare human-macaque differential expression of the human-specific DE genes along the droplet-based pseudotemporal trajectory and the C1-based fetal brain pseudotemporal trajectory.

Single-cell and bulk ATAC-seq data generation. Two-month-old and four-month-old organoids were washed twice with PBS in a Tissue-Tek Cryomold (Sakura), then embedded in 4% low-melting agarose (Sigma) and sliced into 150 µm sections using a vibrating microtome (Ci 7000 smz, Camden Instruments). Slices were placed on microscope slides containing Differentiation medium with vitamin A (Diff+VA) and inspected under a stereomicroscope to dissect cortical regions. Selected regions were washed twice in 500 µL PBS and incubated at 37°C in 500 µL Accutase (Sigma) plus 0.5 µL DNase I (New England Biolabs) for 45 min. Trituration was performed for additional mechanical dissociation. Cells were passed through a 30 µm pre-separation filter (Miltenyi Biotec), washed with Diff+VA medium, and spun down at 300g (Heraeus Megafuge 40R, Thermo Scientific) for 5 min. The cell pellet was resuspended in 200 µL of Diff+VA medium. Cells were viewed under a microscope to ensure a single cell suspension was obtained, and then counted using a Countess Automated Cell Counter (Invitrogen). Single cell suspensions for the early stages of organoid differentiation (iPS cells to neuroepithelium) were obtained as described above.

From the cell suspension, 50,000 cells were used as input for bulk ATAC-seq as described [285]. The remaining cells were diluted to a final concentration of 300 cells/µL and used for microfluidics based single-cell ATAC-seq as described [197]. In brief, cells were mixed with Suspension Reagent (Fludigm) at a 3:2 ratio and loaded onto a primed medium (10-17 µm) integrated microfluidic circuit (Fludigm) for capturing. Cell capture sites were examined under a microscope and noted for containing 0, 1, or multiple cells. Lysis, transposition, and amplification were performed on the Fluidigm C1 platform. DNA from each cell was transferred to an individual well of a 96-well plate and barcoded with unique combinations of 24 adaptor-index i7 and 16 adaptor-index i5 primers [197]. Quantification and library size distribution was assessed on an Agilent 2100 Bioanalyzer using High Sensitivity DNA chips. Excessive primer contamination was removed using SPRIselect (Beckman Coulter Life Sciences) size selection. Up to 192 cells were pooled and sequenced in pairedend, dual-index mode for 50 + 8 + 50 + 8 cycles on one lane of an Illumina HiSeq 2500. A summary of all single-cell experiments can be found in Supplementary Table 1.

Single-cell and bulk ATAC-seq data processing. Base calling was performed using Bustard (Illumina), adaptor trimming with leeHom [261], and demultiplexing with deML [262]. Reads were aligned to hg19 for human, panTro4 for chimp, and rheMac8 for macaque using bowtie2 with a maximum fragment length of 2,000. PCR duplicates were marked and removed using Picard tools (http://broadinstitute. github.io/picard). Samtools [286] was used to retain properly paired reads with mapping quality greater than 30, while reads mapping to the mitochondrial genome, Y chromosome, and blacklisted genomic regions that show excessively high read mapping, several of which correspond to nuclear mitochondrial DNA segments (identified in ref. [197] and the ENCODE Project [287]) were removed. For scATAC-seq, single cell BAM files were merged, excluding data from any capture site with 0 or more than 1 cell, to create an aggregated BAM file. Peaks, which represent regions enriched in mapped pair-end sequences, were called using MACS2 [288] with options nomodel, nolambda, keep-dup all, and call-summits. Peak summits were extended by ±250 bp. In the event of overlapping peaks, the peak with the lowest P value was kept. A single-cell ATAC-seq consensus peak set was obtained by requiring a peak to be accessible in a minimum of 5% of cells. Data visualization was carried out using the Integrative Genomics Viewer (IGV) [289].

Enrichment for validated human VISTA enhancers. We overlapped scATAC-seq peaks detected in human cerebral organoids with positive human VISTA enhancers using bedtools intersect. For each tissue annotated in the VISTA Enhancer Browser, we counted the number enhancers that did or did not overlap a peak. We compared these values to the number of all other tissue elements that did or did not overlap a peak. Fisher's exact tests were performed to determine which tissues' enhancers had a higher likelihood of being represented. The significance values were corrected for multiple testing using the qvalue package in R.

Cell state identification using single-cell ATAC-seq on cerebral organoids and pseudotime estimation. The accessibility at each site in the consensus peak set for every single cell was used to create a count matrix. Cells with fewer than 5,000 read pairs and less than 5% of reads in peaks (fraction of reads in peaks, FRiP) were filtered out from further analyses. chromVAR [278] was used to scan the peaks for transcription factor binding motif occurrences, using a curated collection of 1,765 human motifs from the cisBP database, and to identify significantly variable motifs among cells. In addition to TF binding motifs, peaks were scanned for 7-mers. Cell similarity was visualized in a two-dimensional *t*-SNE plot using the bias-corrected deviations in accessibility for 7-mers.

Each cell's t-SNE coordinates and the consensus peaks were passed to Cicero [290] and the densityPeak algorithm was used to identify two clusters of cells. Statistically significant differences in TF motif accessibility between the two clusters was calculated using chromVAR, and those motifs corresponding to marker transcription factors known to distinguish neural progenitors and neurons was used as the basis for cell state identification. Statistically significant differences in accessibility of additional annotations between the two clusters were used to support cell state identities. These annotations included differentially accessible chromatin peaks identified as being enriched in developing mouse brain radial glial cells or excitatory neurons [291], as well as accessibility in peaks nearby genes showing pseudotime-dependent expression in cortical neural progenitors or cortical neurons identified as part of this study.

We identified differentially accessible (DA) peaks between the two clusters using the command differentialGeneTest in Cicero. A count matrix was generated with featureCounts [292] using the top 250 DA peaks in each cluster. This count matrix was used as input for a diffusion map in order to obtain a pseudotemporal ordering of the cells [293]. Projecting transcription factor binding motif deviation Z-scores on the cells revealed a gradient of known neural progenitor to neuronal markers along the first diffusion map component and we took a cell's rank along this component as its pseudotime value.

DA peaks identified between the two clusters were used as input test regions for GREAT (v.3.0.0) [294] with all accessible organoid peaks serving as background regions. We used the default basal plus extension genomic association rule with its default values. All gene ontology (GO) Biological Process terms and their associated hypergeometric P values were exported. For each term, we plotted the P value obtained using cluster 1 DA peaks and the P value obtained using cluster 2 DA peaks. Terms with $P < 0.05$ were considered enriched. Informative enriched terms were highlighted based on their significance value in one cell state relative to the other, and for small differences between the cell states when highlighting terms enriched in both.

Single-cell ATAC-seq pseudotime estimation for cells in early states of differentiation and cerebral organoids. Similar to the analysis of the cerebral organoids, we used chromVAR to calculate bias-corrected deviations in accessibility for TF motifs and 7-mers for each cell across the differentiation trajectory. Here, we included the scATAC-seq consensus peak sets called in the iPSC, embryoid body, neuroectoderm, and neuroepithelial time points, in addition to the scATAC-seq consensus peak set from the cerebral organoid time point. In the event of overlapping peaks, the peak with strongest signal was retained. Cells with fewer than 5,000 read pairs and less than 5% of reads in peaks were removed from further analyses (Supplementary Table 9). Cell similarity was visualized in a two-dimensional t-SNE plot using the bias-corrected deviations in accessibility for 7-mers.

As the cerebral organoid cells' pseudotimes were previously resolved, we focused on ordering the earlier stages. For this we used Cicero's differentialGeneTest to identify DA peaks among the iPSC, embryoid body, neuroectoderm, and neuroepithelial time points. A count matrix was generated using the top 250 DA peaks in each time point and used as input for a diffusion map. Projecting TF motif deviation Z-scores of the cells revealed a gradient of pluripotent to more differentiated marker TFs along the first three diffusion map components. We fit a principle curve through the map, and used the pluripotent cells as a starting point to guide the curve. The rank of a cell along this curve was used as its pseudotime. We then added the cerebral organoid cells' pseudotime ranks to this earlier stage resolved pseudotime. We used the pheatmap R package to visualize the dynamics of significantly variable TF motifs across pseudotime.

Annotation of accessible chromatin peaks. Peaks were linked to an expressed protein-coding gene using the nearest (maximum distance 1 Mb) transcription start site of the canonical transcript as defined by GENCODE (comprehensive gene annotation, release 19). Promoter regions were defined as 1,000 bp upstream a TSS, and distal regions refer to non-promoter regions. Exon and intron annotations were also obtained from GENCODE (comprehensive gene annotation, release 19). BEDtools [295] was used to annotate peaks for several evolutionary signatures, including: human accelerated regions [296–298]; selective sweeps compared to great apes [299] and archaic humans [300]; SNCs in modern humans that happened since the split with great apes and before or after the split with the ancestor of Neandertals and Denisovans, first identified in ref. [301] and updated for this analysis using the most current 1,000 Genomes Phase 3 allele frequencies, with a global allele frequency $\geq$ 99.5% defined as fixed in all modern humans; small insertions and deletions (up to 5 nucleotides) fixed in modern humans that happened since the split with great apes and before or after the split with the ancestor of Neanderthals and Denisovans [302]; and, human deletions that are highly conserved in mammals [60] (hCONDELs, Supplementary Table 11).

Identification of genomic regions with differential accessibility between human and chimpanzee organoid neural progenitors and neurons. To compare the chromatin accessibility of NPCs and neurons in cerebral organoids between human and chimpanzee and identify putative regulatory regions that may contribute to transcriptome divergence between human and chimpanzee, we applied a likelihood ratio test based on a generalized linear model with binomial error distribution to each regulatory region identified in human and chimpanzee organoids. More specifically, we identified open chromatin regions in human and chimpanzee organoids separately as described above. To compare an equal number of human and chimpanzee regions, we took the top 77,611 peaks (corresponding to the number of human consensus peaks) in each species and performed reciprocal liftOver, requiring a 50% minimum ratio of bases that must remap, in order to identify their orthologous counterparts in the other species. Peaks that successfully lifted over ($\geq$99%) were merged using bedtools and re-named (that is, mergePeak#). Count matrices were generated at these merged peaks in the species own genome, and the matrices were then joined on the common peak name. Considering the higher read coverage in human cells, we subsampled reads in human cells to equalize the medians of total number of reads mapped to the regions of interest in human and chimpanzee. This procedure was applied separately to NPCs and neurons. The resulting count matrices were binarized. We then fitted a generalized linear model for each region across all human and chimpanzee cells, with the accessibility as the response variable and species as the independent variable. Another model with the species variable replaced by a scaling coefficient was also fitted as the null model. The scaling coefficient is fixed to one for human cells and $p_\mathrm{c}/p_\mathrm{h}$ for chimpanzee cells, where p_c and p_h are the average accessibility across all regions and all cells in chimpanzee and human, respectively. We compared the two models and got the P values by using the likelihood ratio test. Regions with BH-corrected $P < 0.01$ were defined as DA regions (Supplementary Table 10). This procedure was applied to NPCs and neurons separately to obtain DA regions in the two cell states.

Functional and evolutionary characterization of genomic regions with differential accessibility. We performed permutations to determine if DA peaks were significantly more likely to overlap a given annotation compared to non-differentially accessible (non-DA) peaks. In more detail, we first resized all peaks to an equal length of 500 bp and calculated the average accessibility of human and chimp cells in the resized DA and non-DA peaks. Peaks were then placed into average accessibility bins of 5% intervals. Given the number of DA peaks in each accessibility bin, the same number of non-DA peaks was chosen at random from the corresponding accessibility bin. The random set of non-DA peaks was then overlapped with the given annotation using bedtools intersect. The random sampling of non-DA peaks and annotation overlap was repeated 2,000 times. For each annotation, we counted the number of times a non-DA peak permutation resulted in a higher overlap than what was observed for DA peaks. This number was divided by the number of permutations to determine significance ($P < 0.05$).

We used fixed SNCs, organoid-specific peaks, and linked differentially expressed (DE) genes as annotations. When overlapping peaks with fixed SNCs, we restricted the analysis to include only regions that passed a stringent genome alignability filter ("map35_100%")[301], in which SNCs could be called. Organoid-specific peaks were defined as peaks detected in 2-month and 4-month old cerebral organoid stages, but not detected in earlier stages of differentiation (pluripotency to neuroepithelial stages). Cell state-specific peaks were those identified as differentially accessible between NPCs and neurons in either human or chimp.

To study putative effects of fixed SNCs on transcription factor (TF) binding in the accessible genomic regions, we used funseq2 [303] to scan and statistically evaluate all possible TF binding motifs created by fixed SNCs in DA peaks. To generate a list of TF motifs lost on the human lineage, we used the human allele as the reference allele and the ancestral allele [301] as the alternative allele. To generate a list of TF motifs gained on the human lineage, we flipped the state of the reference and alternative allele. This allowed us to directly compare the sequence scores of TF motifs gained or lost in humans. We subtracted the sequence score with the alternative allele from the sequence score with reference allele and performed minmax normalization. Human TF motif gains were plotted as positive values, while human TF motif losses were plotted as negative values. The genomic location of SNCs predicted to alter TF motif binding are provided in Supplementary Table 11. The alteration rate for TF motifs gained in humans was calculated by dividing the number of gains in DA peaks by the number of occurrences of that motif when scanning all organoid accessible peaks using chromVAR and the human genome sequence. The alteration rate for TF motifs lost in humans was calculated by dividing the number of losses in DA peaks by the number of occurrences of that motif when scanning all organoid accessible peaks using chromVAR and the chimpanzee genome sequence. The alteration rates of human TF gains and losses were also calculated per TF family, using TF motif family assignments obtained from [304].

We used the macaque cerebral organoid scATAC-seq data to determine species specificity of the peaks identified as differentially accessible between human and chimpanzee (Supplementary Table 10). In brief, we counted read coverage of each accessible region we compared between human and chimp which can lift over to the macaque genome in each macaque cell. Regions failed during liftover were seen as inaccessible in all macaque cells. A random sampling of reads in human and chimpanzee cells was applied to equalize median read coverage in the three species. This procedure was applied 100 times and to the two cell states separately. Accessible probability was then calculated for the two cell states in the three species. In human and chimpanzee, averages across the 100 read-subsampling-based estimation were used. The difference of accessible probability between human and macaque (H - M), and that between chimpanzee and macaque (C - M), was then calculated for each human-chimpanzee DA peak in each cell state. The identified DA was considered as human-specific if its H–M difference is at least four times larger than the C–M difference, while its H–M difference is no less than 2%. Similar criteria were also applied to define chimp-specific DA.

To investigate potential biological processes that may be influenced by DA peaks, we used human-chimp DA peaks for each cell state (NPC or neuron) as input test regions for GREAT (v.3.0.0) [294] with all accessible organoids peaks serving as background regions. This analysis was then carried out the same way as explained above.

Single-nucleus and bulk RNA-seq data generation. Tissue cubes were dissected from frozen post-mortem prefrontal cortex tissue from human, chimpanzee, bonobo and macaque individuals. In total, three healthy adult human (*H. sapiens*), two healthy adult chimpanzee (*Pan troglodytes*), one healthy adult bonobo (*Pan paniscus*) and three healthy adult rhesus macaque (*Macaca mulatta*) brains were used. The human samples were obtained from the Chinese Brain Bank Center (CBBC) in Wuhan, China. For each of these individuals, written informed consent to use human tissues for research was obtained either from the donors themselves or from their next of kin. All subjects were classified as normal by forensic pathologists at the brain bank, and suffered sudden deaths with no prolonged agonal state. One chimpanzee sample was obtained from Biomedical Primate Research Centre, Netherlands; the other chimpanzee sample was obtained from Yerkes National Primate Research Center, USA. The bonobo sample was obtained from Lola ya Bonobo Sacturary, Congo. Tissue was shipped to the Max Planck Institute for Evolutionary Anthropology under the institutional permit for the transport of biological material derived from endangered species (DE216-08, http://cites.org/common/reg/si/e-si-beg.shtml). Rhesus macaque samples were all collected at the Primate Research Center in Goettingen, Germany. All the chimpanzee and bonobo individuals suffered sudden deaths for reasons other than their participation in this study and without any relation to the tissue used. All the macaque individuals were euthanized. The Biomedical Research Ethics Committee of Shanghai Institutes for Biological Sciences reviewed the use and care of the animals in the research project (approval ID: ER-SIBS-260802P).

Dissection was performed on dry ice aiming for cubes with minimal curvature to obtain reproducible slicing results. In brief, the thickness of grey matter at all facets of the cube was measured to obtain a mean grey matter thickness. The mean thickness was divided by 10 to obtain the thickness for each of the segments, whereby each of the segments consisted of several slices at 50 µm thickness. Sectioning was performed in a cryostat (Microm, Thermo Fisher), with slices being alternately immersed in Trizol (Invitrogen) for bulk RNA isolation or transferred to a dry tube (low binding) for single nucleus isolation on dry ice. Segments 11 and 12 were collected as well but were considered being derived from white matter of the cortex. Samples were then stored at −80 °C until further use.

For nuclei isolation from frozen tissue, all following steps were performed on ice with precooled buffers and centrifugation steps were performed at 4 °C. In brief, tissue was spun down, thawed on ice and 1 ml PBSE (PBS (Gibco), 2 mM EDTA (Life Technologies)) was added to the tissue. The tissue slices were incubated at 4 °C on a shaker at 1,500 r.p.m. for a total of 45–60 min with trituration steps in between using 1,000p and 200p to homogenize the tissue. Generally, segments 1–10 were used for single-nucleus experiments. Two segments were pooled to obtain sufficient material for single nucleus isolation, resulting in 5 segments per individual. To reduce batch effects and increase the number of nuclei per experiment, material from three different individuals (originating from human, chimp/bonobo and macaque respectively) was pooled for each segment. After homogenization, solutions were combined in a 5 ml tube and spun down at 900g for 5 min. The pellet was resuspended in 1.5 ml PBSE + 1% NP-40 (BioVision), triturated 20 times using 1,000p and incubated for 7 min incubation on ice. Samples were then spun down at 900g for 5 min and resuspend in 1.5 ml PBSE + 1% BSA (Serva) two times. Samples were then spun down again at 900g for 5 min and resuspended in PBS + 1% BSA. Before sorting, samples were filtered through a 30 µm cell filter (Miltenyi Biotec) and stained using DAPI (1:1,000, BD Pharmingen). Nuclei were sorted in yield sort mode (BD FACS AriaIII and BD FACS Fusion) based on a defined nuclei population by excluding debris using FSC and SSC and by sorting DAPI positive events. Nuclei were sorted in bulk into 96 well plates and spun down 5 min at 600g to enrich for nuclei in the pellet.

For each of the pooled samples, two lanes on a 10x Chromium microfluidic chip were loaded, aiming for the maximum possible number of nuclei to be targeted obtained from the sorting. Single-nucleus experiments were performed using the 10x Genomics Single Cell 3' kit v.2 to encapsulate nuclei along with barcode tagged beads, generate and amplify cDNA and to generate sequencing libraries. Each pooled library was barcoded using i7 barcodes provided by 10x Genomics. cDNA and sequencing library quality and quantity were determined using Agilent's High Sensitivity DNA Assay. Libraries were pooled and sequenced in 150-bp paired-end mode on Illumina's NovaSeq platform as provided in Supplementary Table 1.

RNA isolation for bulk-RNA Seq was performed using the Direct-zol 96 RNA kit (Zymo Research) and was quantified using Agilent's Bioanalyzer RNA 6000 Nano and Pico kit. Libraries were prepared using the NEBNext Ultra Low RNA Library Prep Kit (New England Biolabs). Library quantification was performed using Agilent's Bioanalyzer DNA 1000 chip kit. All bulk RNA Seq libraries were pooled at equal ratios and sequenced on one lane of an Illumina NovaSeq platform in 150 bp paired-end mode.

Processing of single-nucleus and bulk RNA-seq data from human, chimpanzee and macaque adult brains. Single-nucleus libraries were demultiplexed based on their i7 index sequences using 10x Cell Ranger (v.2.1). Mapping to the humanchimp- macaque consensus genome and generation of count matrices was then performed using the same Cell Ranger, with the GENCODE v.27 human annotation provided. Nuclei were assigned to species based on species specific sites using a two-step approach by separating all great ape from macaque nuclei first and subsequently assigning nuclei to either human or chimp/bonobo. Nuclei with a support of less than 80% for either of the groups were removed from further anal ysis. Moreover, nuclei with less than 200 and more than 6,000 genes detected, so as those with more than 5% detected transcripts being transcribed from mitochondria, were removed from further analyses.

The full snRNA-seq dataset including all species was further analysed using Seurat (v.3) (Supplementary Table 13). Single-nucleus expression values were normalized and highly variable genes were identified using a variance stabilizing function to detect the top 2,000 variable genes (Supplementary Table 12). Data were then integrated by finding corresponding anchors between the species using 30 dimensions. Scaling and PCA were performed using the integrated data. The top-20 PCs were used to identify neighbours of cells and clusters and to visualize the clustering using *t*-SNE embedding. Cluster identities were assigned using unbiased identification using cluster markers by running Seurat's FindAllMarkers function (Wilcoxon test, min.pct = 0.25, min logFC = 0.25) using non-integrated expression values, known marker genes reported elsewhere [215, 221] and by cell type prediction using Seurat's

TransferData function to anchor to the published Dropseq based human adult frontal cortex snRNA-seq data [221]. Two potential doublet clusters (c11, c19) were excluded from further analysis. For analysis of the major cell classes (excitatory neurons, inhibitory neurons, astrocytes, oligodendrocytes, oligodendrocyte precursor cells, microglia, endothelial cells) subtype clusters were combined and cell type markers recalculated using Seurat's FindAllMarkers function (Wilcoxon test, min.pct = 0.25, min logFC = 0.25) using non-integrated expression values (Supplementary Table 14).

Since nuclei of the three species have significantly different transcriptome coverage, pseudonuclei were constructed for more robust transcriptome measurement, as well as for more fair and efficient comparison, using a similar procedure as described above to generate pseudocells, under the constraint of merging only nuclei from the same segment of the same sample and grouped in the same cell cluster. The probabilities of nuclei selected as pseudo-nuclei seed were 1/13 for human, 1/8 for chimpanzee and 1/10 for macaque.

Reads of the bulk RNA-seq samples were mapped to the human-chimpanzee- macaque consensus genome using STAR (v.2.6.1d). The Python utility hiseq-count was used to count the numbers of uniquely mapped reads of genes annotated in GENCODE v.27 human annotations. DESeq2 was used for normalization and retrieving FPKM as the expression level measurement.

To determine the laminar origin of each segment, genes with segment-dependent expression were first screened for each cortical cube. In brief, an ANCOVA analysis was applied to compare two models: the natural spline (df = 6) linear model with $\log_{10}$-transformed FPKM as the response and the segment order as the variable; the null model of expression values without any relationship with segments. For each of the resulted gene, its enriched segments in the cube were identified, as the segments with the gene's expression at least one standard deviation higher than the mean across segments. Genes with enriched expression at each segment were then overlapped with the layer markers identified in [141]. Segments with enriched genes significantly overlapping with markers of only one layer were seen as pure-layer original, others were seen as mixture of multiple layers. For each mixture segment, a quadratic-programming-based transcriptome deconvolution [305] was applied to determine the relative contribution of the enriched layers. A layer index was then obtained for each segment, as the average layers weighted by contributed proportion of the enriched layers.

Estimation of cell type distribution across cortical layers and gene expression patterns in neurons across cortical layers. To estimate the cell type composition of each layer, nuclei from each sample were randomly assign to one layer, based on the layer mixture proportions estimated above. The proportion of each of the six major cell classes: excitatory neurons, inhibitory neurons, astrocytes, oligodendrocytes, oligodendrocyte precursor cells (OPCs), microglia and endothelial cells, was then calculated for nuclei assigned to each layer in human. This procedure was repeated 100 times, with the resulted average as the final estimation. The laminar distribution of each cell cluster was also estimated based on the described procedure. In addition, a subsampling procedure with replaceable manner of the same number of nuclei ($n = 200$) from each layer was further applied to each of the 100 nuclei layer random assignment to control differences on the detected nuclei number of each layer.

To get more precise estimation of layer origins on the nuclei level for excitatory and inhibitory neurons, both of which show a distinct layer distribution pattern across different subtypes, we trained an elastic net linear regression model ($\alpha = 0.5$) on excitatory and inhibitory neurons separately, with the sample layer indices as the training response and expression levels of the highly variable genes as the variables. To enhance model robustness, pseudo-nuclei from all the three species together were used for model trainings. The trained models were then applied to the excitatory and inhibitory pseudo-nuclei again. The predicted layer indices were used as the estimated relative laminar location of the pseudo-nuclei. The projection of the predicted layer indices to layers were done by averaging expression patterns of markers of different layers [141].

Differential expression analysis between human and chimpanzee cell types in adult brains and determination of their species-specificity. Due to the sparse nature of the snRNA-seq data and the unequal coverage of nuclei from different species, commonly used statistical test for differential expression analysis (for example, Wilcoxon's rank-sum test) failed to provide reliable estimation of DE, even with the state-of-art VST normalization methods [306]. As detection rates of genes are correlated with their expression levels [306], we therefore compared gene expression levels of the same cell type in human and chimpanzee by comparing their detection rates, using a GLM-ANCOVA analysis similar to the one described above to identify genomic regions with differential accessibility. In brief, the pseudo-nuclei expression matrix was binarized. A binomial GLM model was trained for each gene, with its detection as the response variable and species of pseudo-nuclei as the independent variable. This model was compared to the null model with the species variable replaced by a scaling coefficient. The scaling coefficient is fixed to one for human pseudo-nuclei and p_c/p_h for chimpanzee pseudo-nuclei, where p_c and p_h are the average detected gene numbers across pseudo-nuclei involved in the test in chimpanzee and human, respectively.

While the described DE test was applied to four cell classes with sufficient numbers of pseudo-nuclei: excitatory neurons, inhibitory neurons, astrocytes and oligodendrocytes, the heterogeneity within the two neuron types, as well as their uneven distributions in human and chimpanzee, needed to be considered. A subsampling procedure with replaceable manner was therefore applied. In every subsampling, an equal number of pseudo-nuclei ($n = 200$) from each species were sampled, with pseudonuclei in clusters annotated as the cell class of interest sharing equal probability being selected. The described DE test was then applied to the sampled nuclei of this cell class. This subsampling procedure was repeated for 100 times, and DE genes of each cell class were defined as genes with significant DE (BH-corrected $P < 0.005$) in at least 80 times of the subsampling. Additional filtering was then applied, requiring the same direction of human-chimpanzee difference on detection rates and VST-normalized expression values.

Macaque pseudo-nuclei were then introduced to investigate species specificity of the identified DE. Similar procedure sampling the same number of pseudo-nuclei from clusters annotated to be the same cell class was repeated 100 times to the macaque pseudo-nuclei. For each sampling, average VST-normalized expression values were calculated for each cell class in human, chimpanzee and macaque, with which differences between human and macaque (dHM), as well as between chimpanzee and macaque (d_{CM}), were calculated. The identified human-chimpanzee DE was defined as human-specific if $|d_{\mathrm{HM}}| > 4 \times |d_{\mathrm{HC}}|$. Genes with chimpanzee-specific DE were identified in the same way (Supplementary Table 15).

Statistics and reproducibility. In the box plots of Figs. 1g, 2d, boxes represent IQR, whiskers represent minimum and maximum with outliers removed. In the box plot of Fig. 4c, boxes represent IQR, and whiskers represent 1.5 x IQR. In the bean plots of Extended Data Fig. 3g, shapes of beans represent Pearson correlation distributions, dashed lines represent medians of groups.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Sequence data that support the findings of this study have been deposited in ArrayExpress with the accession codes E-MTAB-7552 (single-cell RNA-seq data based on 10x Genomics), E-MTAB-8234 (single-cell RNA-seq data based on Fluidigm C1/Smart-Seq2), E-MTAB-8089 (single-cell ATAC-seq of human organoids), E-MTAB-8043 (single-cell ATAC-seq of chimpanzee organoids), E-MTAB-8083 (single-cell ATAC-seq of bonobo organoids), E-MTAB-8087 (single-cell ATAC-seq of macaque organoids), E-MTAB-8228 (the bulk ATAC-seq data), E-MTAB-8230 (snRNA-seq data of the adult brain samples) and E-MTAB-8231 (bulk RNA-seq data). The expression data are also available for exploration in scApeX via the link https://bioinf.eva.mpg.de/shiny/sample-apps/scApeX/.

Code availability

The computational code used in this study is available at GitHub (https://github.com/quadbiolab/primate_cerebral_organoids) or upon request.

Acknowledgements We thank D. Wollny, A. Brazovskaya, K. Köhler, T. Schaffer, B. Schellbach, A. Weihmann, R. Schultz, I. Bünger, M. Dannemann, R. Snabel, B. Vernot, W. Hevers, M. Schörnig, J. Kelso and K. Sekine for their help with this project; A. Fischer,

M. Halbwax, K. Köhler and the Tchimpounga Sanctuary for support with generating the JoC iPSC line; L. Berninger and J. Peters for contributing Sc102a1 and SandraA organoids; B. B. Lake for sharing nuclei isolation protocols; and A. Fischer, M. Halbwax and the Lola ya Bonobo Sanctuary in Congo for contributing bonobo tissue for single-nucleus experiments. Karyotyping was supported by the Stem Cell Engineering Facility of CMCB at Technische Universität Dresden. Sorting was in part performed at the CUDZ at the Veterinary Medicine Faculty at the University of Leipzig. The silhouette images in the figures are credited to istockphoto.com for human (cole matt) and chimpanzee (A-digit), and flaticon.com for macaque (Freepik). This work was supported by the Max Planck Society (B.T.), Chan–Zuckerberg Initiative (B.T. and J.G.C.), European Research Council (Anthropoid, J.G.C.; Organomics; B.T.) and the NOMIS Foundation (S.P.). S.K. was supported by the Boehringer Ingelheim Fonds.

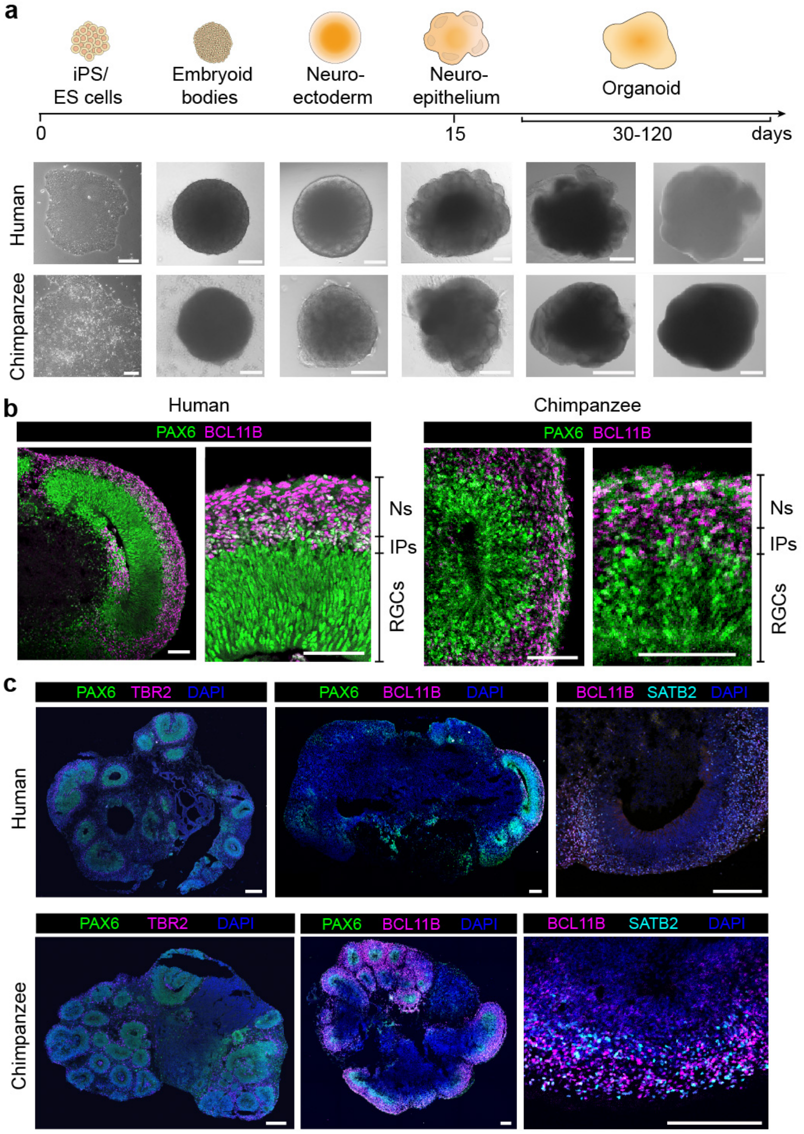
Author contributions S.K. and M.J.B. grew organoids with assistance from A.W., L.S. and M.H. S.K. performed scRNA-seq and snRNA-seq with assistance from M.S. M.J.B. performed scATAC-seq. Z.H., M.J.B. and S.K. analysed the data. F.S.C. and M.H. performed immunohistochemical stainings. J.S.F. compared organoid scRNA-seq data to mouse voxel maps. P.G. dissected and sliced tissue for snRNA-seq. D.H. and Z.Q. performed bulk RNA-seq of adult tissue. S.K., M.J.B., Z.H., B.T. and J.G.C. designed the study and wrote the manuscript with support from P.K., W.B.H. and S.P.

Competing interests The authors declare no competing interests.

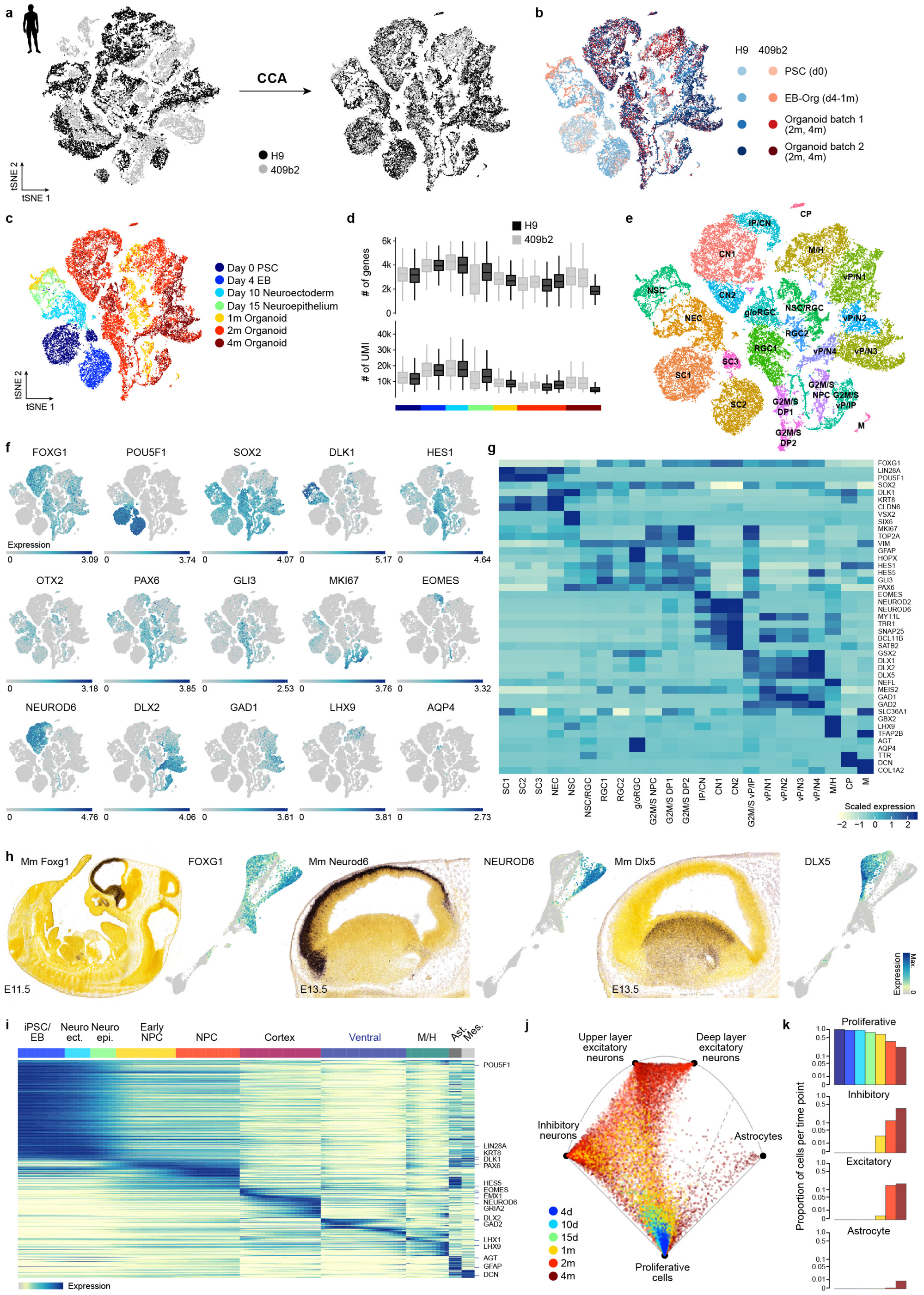
Additional information

Supplementary information is available for this paper at https://doi.org/ 10.1038/s41586-019-1654-9.
Correspondence and requests for materials should be addressed to Z.H., B.T. or J.G.C.
Peer review information Nature thanks Christopher Walsh and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/ reprints.

Extended data figures

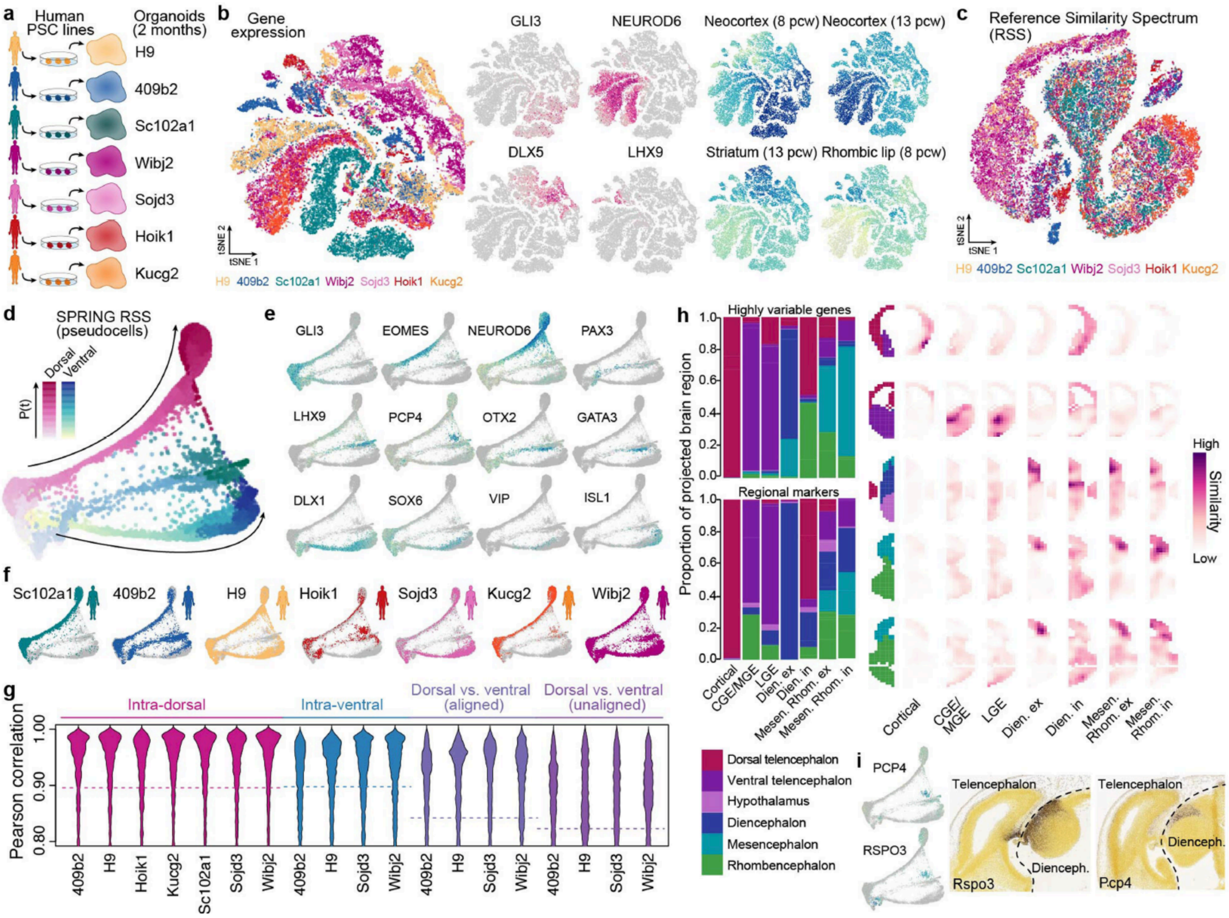


Extended Data Fig. 1: Differentiation and immunohistochemical characterization of human and chimpanzee cerebral organoids. **a**, Phase contrast (PSC to neuroepithelium; scale bar, 200 µm; H9 for human, SandraA for chimpanzee) and bright-field images (organoid; scale bars, 1 mm; H9 and Wibj2 for human, JoC and SandraA for chimpanzee) showing examples of different stages of organoid development for human and chimpanzee. **b**, Immunohistochemical staining for *PAX6* (green) and *BCL11B* (also known as *CTIP2*) (pink) of a 63-day human organoid from iPSC line 409b2 (left) and a 63-day chimpanzee organoid from iPSC line SandraA (right), with a magnified view into a cortical-like region (scale bars, 100 µm). **c**, Immunohistochemical staining of human (top left, Sc102a1, 50-day; top middle, 409b2, 63-day, the same organoid as the human organoid in b) and chimpanzee (bottom left, SandraA, 50-day; bottom middle, SandraA, 63-day) organoids (scale bars, 200 µm) for progenitor (*PAX6*) intermediate progenitor (*TBR2*) and deep-layer neurons (*BCL11B*) in whole organoids. Staining for deep-layer (*BCL11B*) and upper-layer neuron (*SATB2*) markers for human (top right, Sc102a1, 45 day) and chimpanzee (bottom right, SandraA, 63 day) organoids.

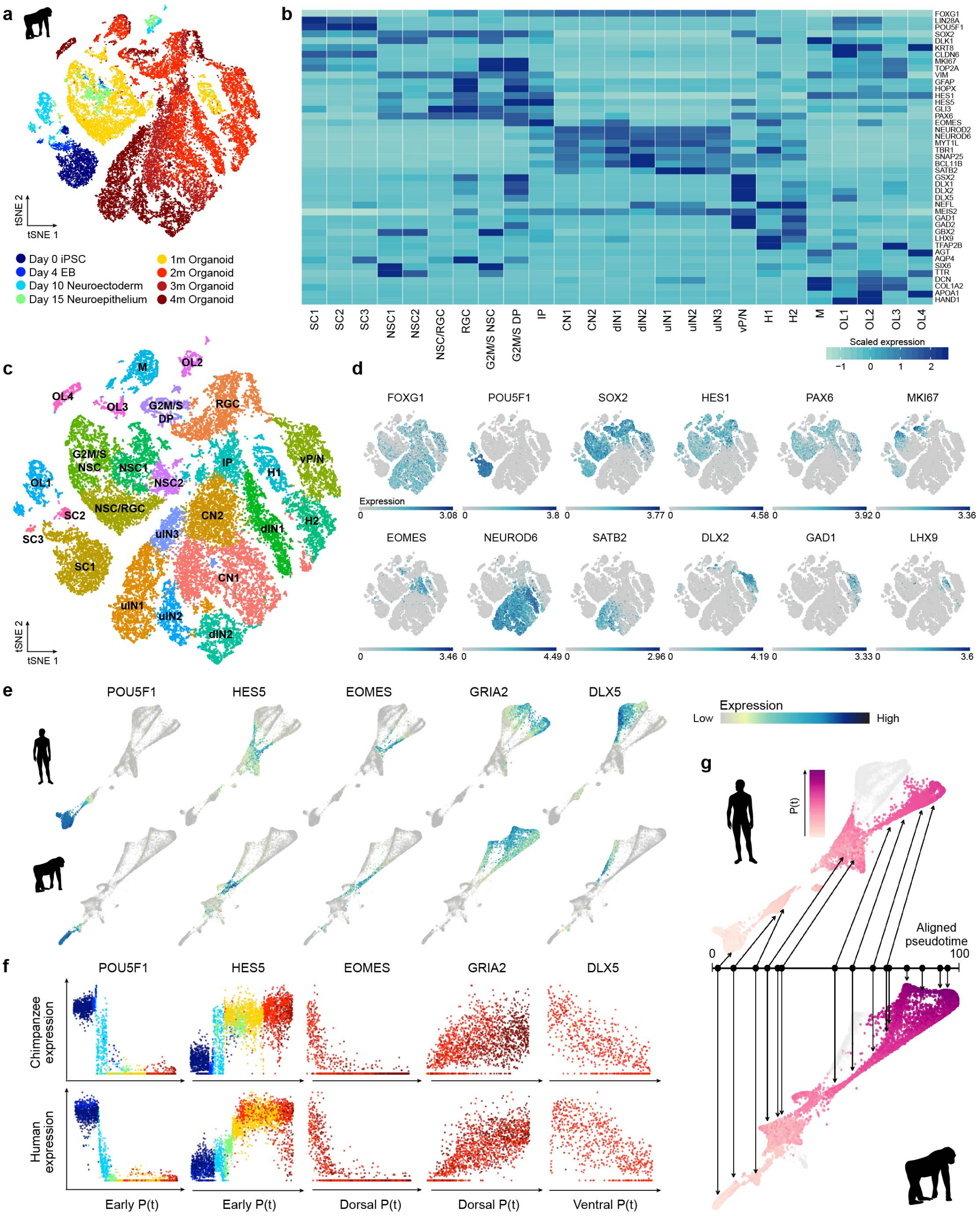


Extended Data Fig. 2: Heterogeneity analysis during human cerebral organoid development from pluripotency. **a**, Cells from different human cell lines (23,226 cells from H9 and 20,272 cells from 409b2) were integrated using canonical correlation analysis and visualized using *t*-SNE. **b**, *t*-SNE colour coded on the basis of cell line and batch. **c**, *t*-SNE coloured on the basis of time point. Heterogeneity analysis was performed on combined cells from day 0 of differentiation to 4-month-old organoids for iPSC and ESC-derived cells. **d**, Distribution of number of genes and UMIs for different time points and cell lines. **e**, Clustering was performed using the top-20 PCs as input for *t*-SNE and cluster names were assigned on the basis of expression of cluster marker genes and known marker genes. SC, stem cells; NEC, neuroectoderm-like cells; NSC, neural stem cells; g/oRGC, gliogenic/outer radial glia cells; G2M/S NPC, neural progenitor cells in G2M/S phase; G2M/S DP, dorsal progenitor cells in G2M/S phase; CN, cortical neurons; G2M/S vP, ventral progenitors in G2M/S phase; M/H, midbrain/hindbrain; CP, choroid plexus; M, mesenchymal-like cells. **f**, *t*-SNE plot coloured by expression level of selected marker genes on the basis of non-integrated expression values.

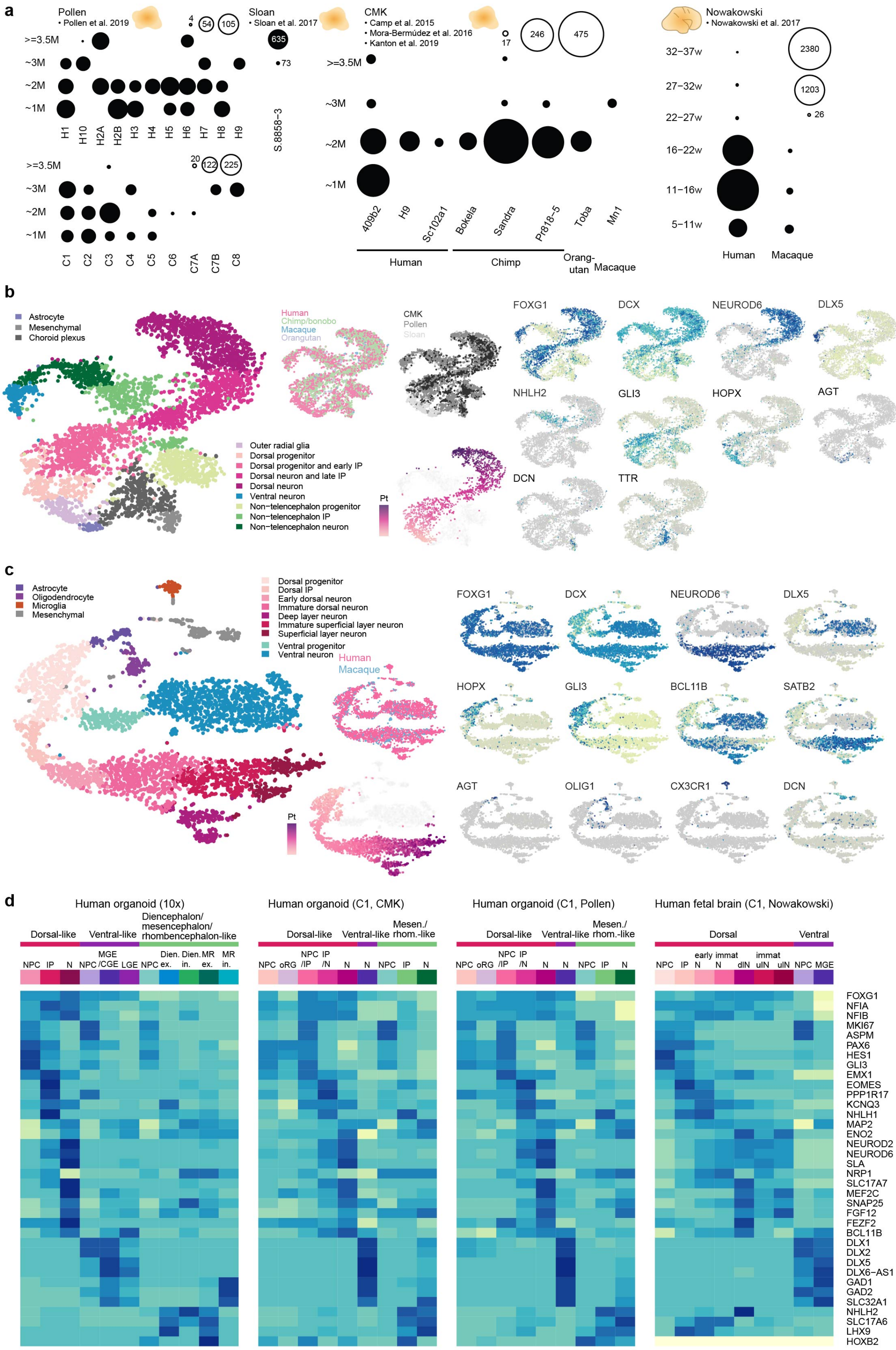
g, Heat map showing averaged cluster expression for representative marker genes for clusters ordered according to their differentiation time from early to later stages and regional identity from dorsal to ventral forebrain and non-forebrain cells. **h**, In situ hybridization images from the Allen Developing Mouse Brain Atlas (available from https://developingmouse.brain-map.org/) showing expression of *Foxg1*, *Neurod6* and *Dlx5* in the mouse developing forebrain and human whole-trajectory SPRING plots coloured by the corresponding genes. **i**, Pseudotemporal expression of example genes marking different stages of development over the whole human cerebral organoid developmental trajectory. **j**, Umbrella plot showing the similarity of each organoid cell to a cell 'prototype' generated from a reference scRNA-seq cell atlas of the human fetal cortex19. **k**, Plots show the proportion of organoid cells per time point that match a reference prototype.



Extended Data Fig. 3: Analysis of human cerebral organoid single-cell transcriptomes from seven individuals. **a**, scRNA-seq was performed on two-month-old cerebral organoids from one human ESC and six iPSC lines. **b**, All data (49,153 cells) were combined and cell heterogeneity was assessed using *t*-SNE with the top 20 PCs as the input. Cells are also coloured by marker gene expression and RSS. **c**, *t*-SNE plot with RSS against Brainspan fetal reference data as the input (RSS-*t*-SNE), coloured by cell lines. Cells from different lines are well integrated. **d**, SPRING plot of two-month-old human organoid pseudocells (9,650), coloured by neuronal trajectory branches and pseudotimes. **e**, SPRING plot of twomonth old human organoid cells, coloured by marker gene expression. **f**, SPRING plots coloured by cell line show contributions of each line to different branches of the trajectory. **g**, Correlations of expression trajectories of genes with pseudotime-dependent expression patterns between cortical cells from each line to the others (pink), ventral cells from each line to others (blue), and cortical and ventral cells from the same lines after or before aligning the cortical and ventral pseudotimes (purple). **h**, Spatial location inference of neuron subtypes in human cerebral organoids. Left, bar plots show proportion of cells of each cell type that show highest gene expression pattern similarity to the average expression patterns in different structures, on the basis of the processed in situ hybridization image data (E13.5) provided in the Developing Mouse Brain database of Allen Brain Atlas (available from https://developingmouse.brain-map.org/). Expression similarity was calculated based on highly variable genes of the scRNA-seq data (top) or regional markers defined with the in situ hybridization data (bottom left). Right, correlation patterns of average regional marker gene expression of each neuron subtype to voxels in five example sections (E13.5), as well as the structural annotation of the sections. **i**, Expression of two marker genes of diencephalon inhibitory neurons (*PCP4* and *RSPO3*) in the SPRING embeddings, and their spatial expression patterns in E13.5 mouse brain (data from Allen Brain Atlas, available from https://developingmouse.brain-map.org/).

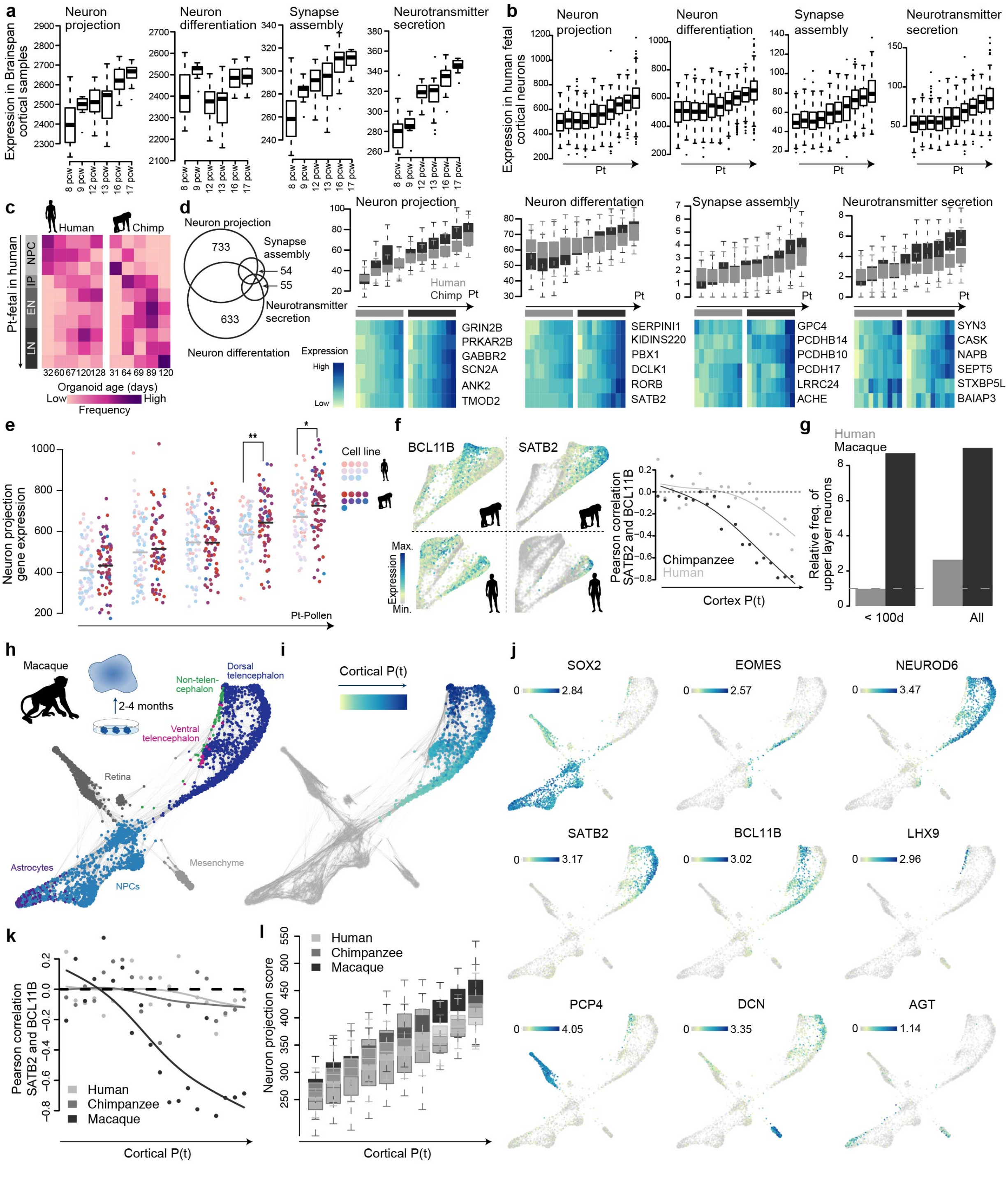


Extended Data Fig. 4: Heterogeneity analysis during chimp cerebral organoid development from pluripotency. **a**, Heterogeneity analysis for iPSC-derived chimpanzee cells (36,884) from day 0 of differentiation to 4 months of organoid development for one cell line (SandraA). **b**, Heat map visualizing averaged cluster expression for marker genes with columns ordered based on differentiation progress from early-to-late time points and regional identity sorted from dorsal to ventral forebrain to nonforebrain cells and non-ectodermal-derived cells. **c**, Cluster identification and *t*-SNE using the top-15 PCs for clustering. Cluster assignment was on the basis of cluster markers as well as expression patterns of known marker genes. SC, stem cells; G2M/S DP, dorsal progenitors in G2M/S phase; dIN, deep-layer neuron; uIN, upper-layer neurons; vP/N, ventral progenitor/neuron; M – mesenchymal-like cells; OL, off-lineage cells. **d**, *t*-SNE plots coloured on the basis of gene expression of representative marker genes used to assign cluster identities. **e**, SPRING plots of whole developmental trajectory for human and chimpanzee coloured by marker genes. **f**, Pseudotemporal gene expression patterns showing marker genes for early, dorsal and ventral branches for human and chimpanzee. **g**, Schematic showing alignment of human and chimpanzee pseudotimes after combining pseudocells from the early stages and the dorsal forebrain lineage. The later chimpanzee pseudotime points fail to align with human pseudocells.



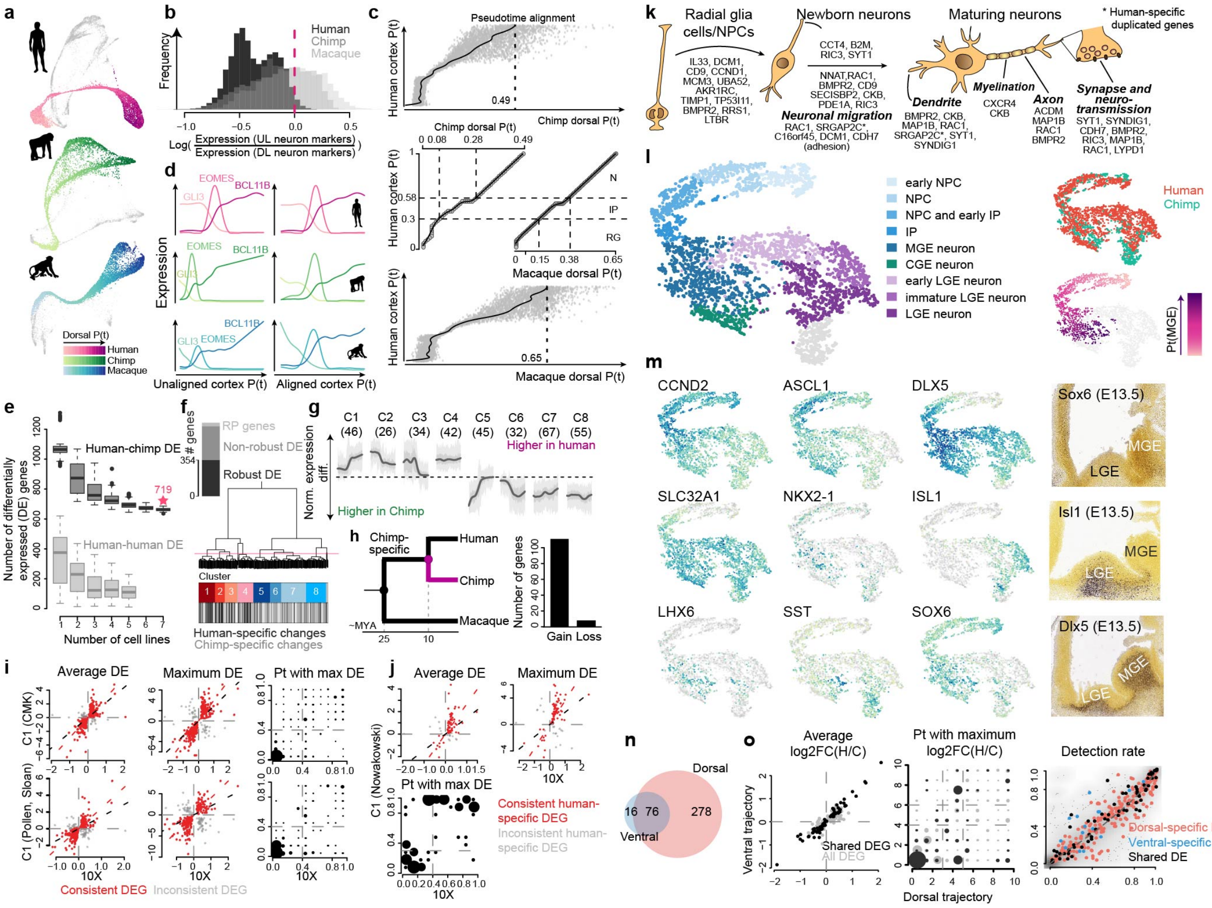
Extended Data Fig. 5: Analysis of cell type heterogeneity of cerebral organoids and fetal cortical tissues based on scRNA-seq data from Fluidigm C1. a, Overview of the Fluidigm C1 scRNA-seq data.

Each dot represents a cerebral organoid or fetal brain sample from one cell line or species at a certain age, with its size showing the number of cells measured. The left panel shows organoid sample information as published in Pollen et al. (2019) [20] (excluding redundant cells from Camp et al. (2015) [114] and Mora-Bermudez et al. (2016) [222]), including the data initially published in Sloan et al. (2017) [133]The middle panel shows organoid sample information generated in Camp et al. 2015 [114], Mora-Bermudez et al. 2016 [222] and in this study. The right panel shows fetal prefrontal cortex sample information reported in Nowakowski et al. 2017 [219]. **b**, All cerebral organoid data (5,838 cells) were combined and cell heterogeneity was assessed using *t*-SNE with the RSS profiles to the fetal Brainspan data as the input. Cells are coloured by cell type or cluster, species, institutions generating the data, dorsal trajectory pseudotimes and marker gene expression. **c**, *t*-SNE plots for all fetal brain data (5,080 cells) to assess cell heterogeneity, with the RSS profiles to the fetal Brainspan references as the input. Cells are coloured by cell type or cluster, species, dorsal excitatory neuron trajectory pseudotimes and marker gene expression. **d**, Heat map showing marker gene expression patterns across different cell types in the droplet-based organoid scRNA-seq data generated in this manuscript and the above described C1-based scRNA-seq data.



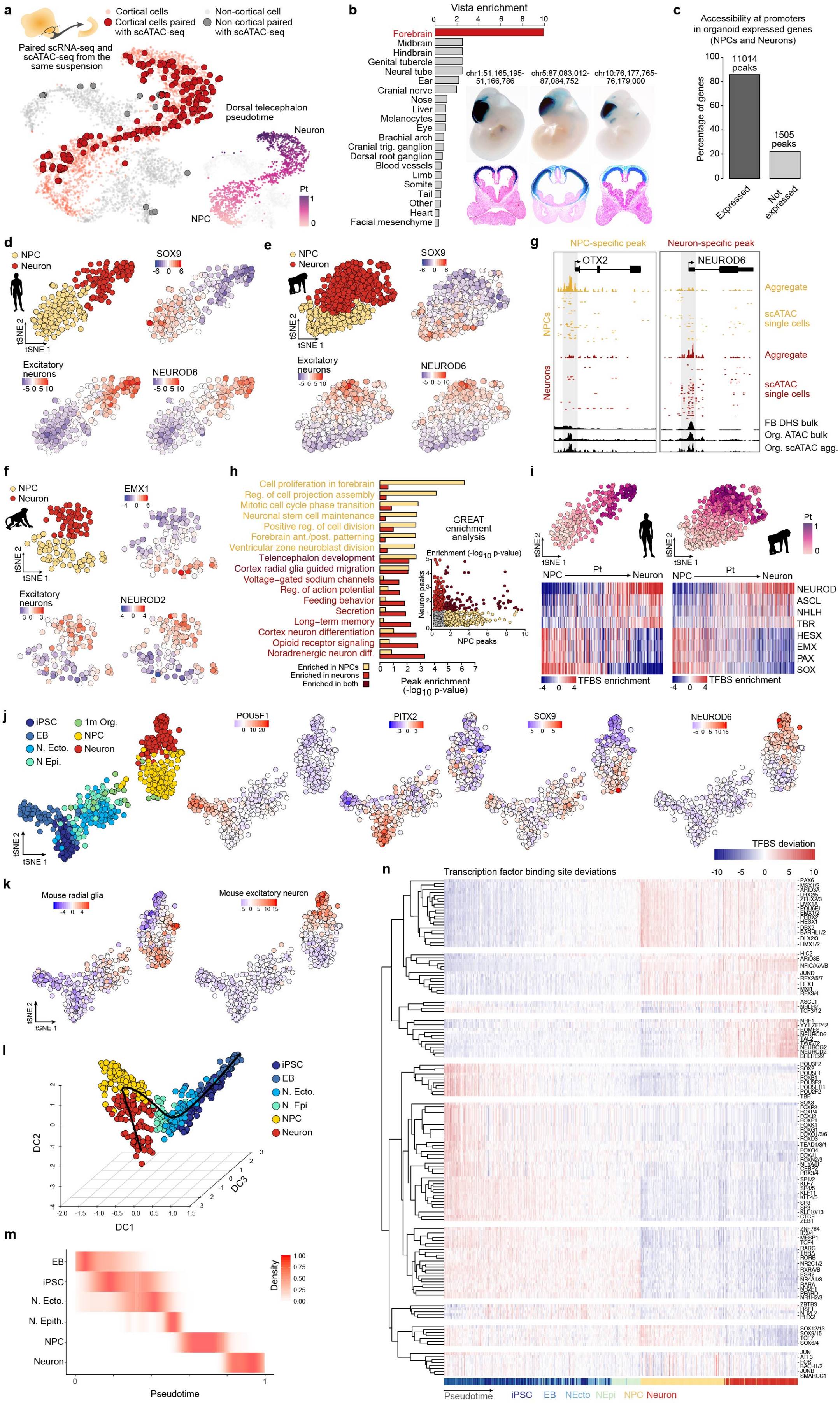
Extended Data Fig. 6: Analysis of neuron maturation timing difference in human, chimpanzee and macaque cerebral organoids. **a**, Box plots (interquartile range with minimum and maximum, outliers removed) showing sum expression levels (in reads per kilobase of transcript per million (RPKM)) of genes with GO annotation neuron projection (1,487 genes), neuron differentiation (1,367 genes), synapse assembly (168 genes) and neurotransmitter secretion (169 genes) in bulk RNA-seq data from Brainspan fetal cortical samples from 8 PCW to 17 PCW. **b**, Box plots showing sum expression levels of the same gene lists in fetal human dorsal excitatory neurons along the estimated developmental pseudotimes (Nowakowski et al. (2017) dataset [219]). **c**, Projection of human and chimpanzee organoid cells to human fetal brain data reveals higher similarity of chimpanzee organoid cells to later stages of development compared to human organoid cells. **d**, Box plots showing sum expression levels of genes with specific annotation to only one of the four GO terms in human and chimpanzee pseudocells (1,791 human and 4,304 chimp) along the cortical pseudotimes.

Heat maps showing expression of example genes from the GO terms for human and chimp along pseudotime bins. The Venn diagram on the left shows the overlap of genes related to the four GO terms. **e**, Distribution of neuron projection scores of human and chimpanzee cortical cells (388 human and 355 chimp) reported in Pollen et al. (2019) along the cortical pseudotimes [20]. Each dot represents one cell, and is coloured by the organoid cell line. Light colours represent human cell lines and dark colours represent chimpanzee ones. Two-sided Wilcoxon's rank-sum test $*P = 0.013$ and $**P = 0.004$). **f**, Observed timing difference of upper-deeper layer specification in human and chimpanzee cerebral organoids from 10x Genomics data generated in this study. The left panel shows expression of cortical deep (*BCL11B*, left) and upper (*SATB2*, right) layer marker genes projected onto the chimpanzee (top) and human (bottom) SPRING plot. *BCL11B* and *SATB2* become anti-correlated in their pseudotemporal expression profile in both human and chimpanzee (right), while the onset of anticorrelation happens earlier in chimpanzee than in human. **g**, Abundance of upper layer neurons relative to deeper layer neurons in human and macaque fetal prefrontal cortex samples [219] in Nowakowski et al. (2017) grouped by early time points ($<$100 days old) or all time points combined. **h**, scRNA-seq was performed on two-to-four month cerebral organoids from a macaque iPSC line. The SPRING plot of pseudocells (2,913) was constructed with the top-20 PCs as the input. The heterogeneity analysis suggests multiple cell types in the macaque organoids, including cortical neurons, NPCs, astrocytes and other cell types such as retina and mesenchyme-like cells. **i**, SPRING plot coloured by pseudotimes of cortical pseudocells, which are the pseudocells' quantiles of diffusion component (DC) 1 of the cortical pseudocells diffusion map. **j**, SPRING plot coloured by marker gene expression. **k**, The onset of anti-correlation between *SATB2* and *BCL11B* occurs earlier along the macaque pseudotime (1,107 pseudocells), relative to human (1,118 pseudocells) and chimpanzee (1,645 pseudocells), when focusing on the two-month cerebral organoids. **l**, Box plots (box shows interquartile range and whiskers show 1.5 x IQR) showing the neuron projection scores in human.



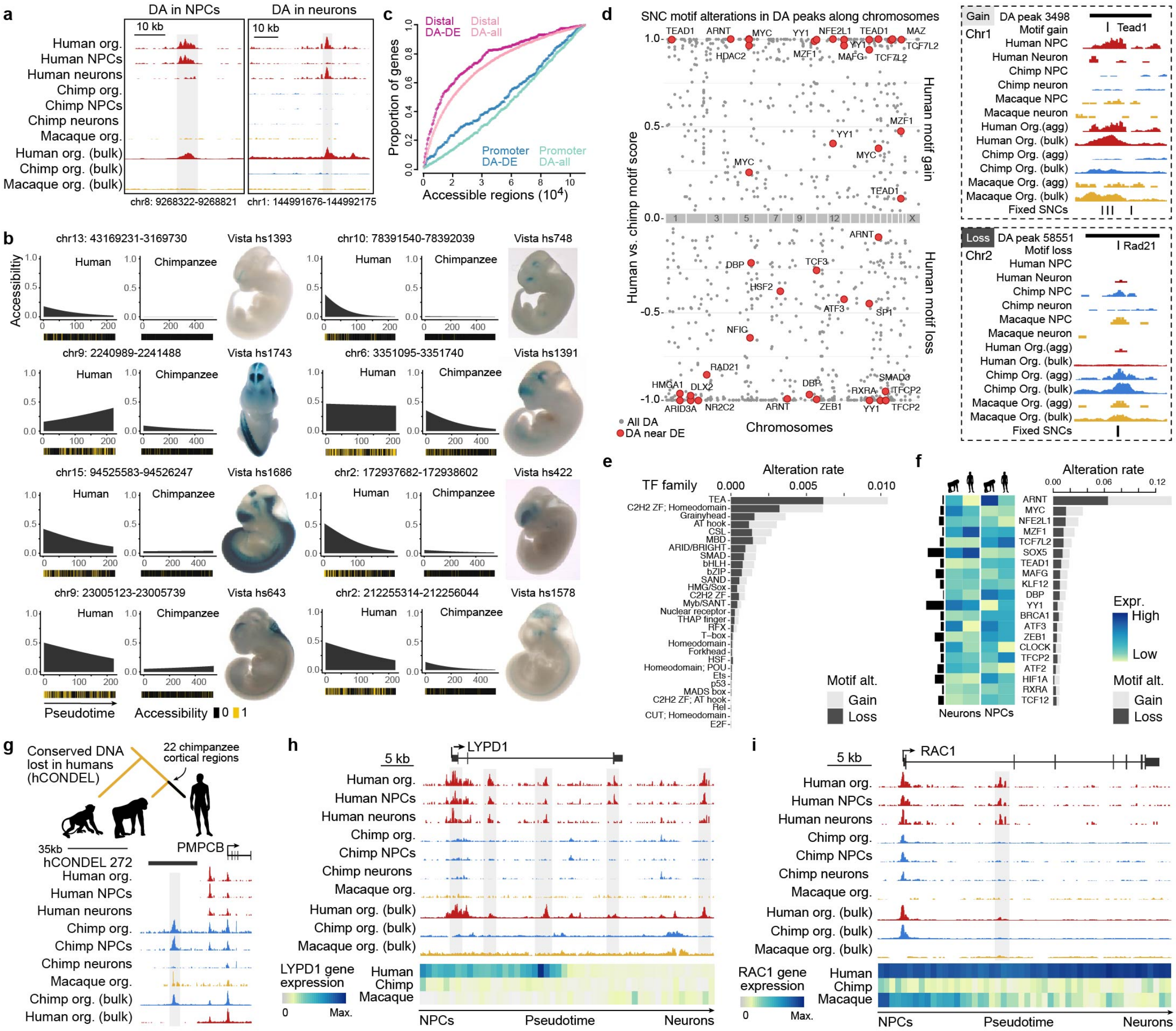
Extended Data Fig. 7: Pseudotime alignment between primates and differential expression between human and chimpanzee. **a**, SPRING plots visualizing the *k*NN networks of human (10,063) and chimpanzee (5,612) pseudocells, and macaque cells (6,580), which represent NPCs and neurons of different brain regions. Cortical NPCs and neurons are coloured by their pseudotimes. **b**, Ratios of upper layer (UL) to deeper layer (DL) neuron marker expression in human (black), chimpanzee (dark grey) and macaque (light grey) organoids. The dashed line indicates the cut-off applied to human pseudocells to filter out those representing UL neurons. **c**, Truncated dynamic time warping (DTW)- based alignment was applied to align human, chimpanzee and macaque cortical pseudotime courses. Two support vector regression models were trained to predict chimpanzee (top) and macaque (bottom) pseudotimes of human pseudocells. A constrained B-splines regression model was fitted to determine the trimming point at the chimpanzee and macaque pseudotime courses, respectively. An end-to-end DTW-based alignment was applied to the human pseudotime course to the trimmed chimpanzee and macaque pseudotime courses for the final alignments (middle). **d**, Pseudotemporal expression profiles of *GLI3*, *EOMES* and *BCL11B* along the human, chimpanzee and macaque cortical pseudotimes, before (left) and after (right) the pseudotime alignment procedures. **e**, Robustness and false-positive rate of differential pseudotemporal expression between human and chimpanzee based on the number of cell lines involved in the analysis with constrained replaceable pseudocell subsampling. In each subsampling, pseudocells representing cells from a certain number of human lines were sampled in a replaceable manner to recapitulate pseudocell distribution along pseudotime course of the chimpanzee pseudocells. Differential expression analysis was applied to compare all chimpanzee pseudocells and the sampled human pseudocells to estimate robustness to cell line numbers (dark grey boxes), and to compare sampled human pseudocells to human pseudocells from another two lines sampled with the same procedure to estimate false-positive rate (light grey boxes). In box plots, boxes represent 100 times of subsampling IQR, the line represents 1.5 x IQR and dots represent outliers. **f**, Robustly detected human–chimpanzee differentially expressed genes (robust DE genes) are defined as the non-ribosomal genes which were detected as DE in at least 80% of the subsampling-based human–chimpanzee DE analysis using any number of human lines (black).

The dendrogram shows the hierarchical clustering of robust DE genes, based on their human– chimpanzee pseudotemporal DE patterns along the aligned pseudotimes of cortical organoid pseudocells, resulting in eight clusters of robust DE genes. **g**, Pseudotemporal differential expression patterns between human and chimpanzee (without including macaque cells) of the eight clusters of genes along the pseudotimes of cortical organoid pseudocells with 50% and 95% confidence intervals shown in dark and light grey, respectively. Numbers of genes in each cluster are shown in parenthesis. **h**, Number of differentially expressed genes in chimpanzee versus human and macaque comparison grouped by gain or loss of expression in chimpanzees. A gain of expression specifically in chimpanzees is more likely than a loss of expression pattern conserved in the other primates. **i**, Comparison of the reported human–chimpanzee pseudotemporal differential expression based on 10x Genomics data with the Fluidigm C1-based scRNA-seq data of human and chimpanzee cerebral organoids. The two rows show the results based on C1 data generated in this manuscript and combined with data from refs [20, 114, 222]. The first two columns show estimated human– chimpanzee differential expression directionality and magnitude in the reported droplet-based scRNA-seq data and the C1-based measurement, with the first column presenting the generalized differential expression along the whole cortical pseudotimes, and the second column presenting the maximum differential expression along the pseudotimes. The red dots represent consistently differentially expressed genes, which have consistent differential expression directionalities in the two datasets. The right panel shows pseudotime intervals with the largest human–chimpanzee differential expression in the two datasets in comparison to the consistent differentially expressed genes. Dot sizes represent frequencies. **j**, Comparison of the estimated human–macaque differential expression directionality and magnitude of the human-specific differentially expressed genes using human and macaque fetal prefrontal cortex scRNA-seq data [20, 222]. **k**, Functional annotations of genes with humanspecific expression patterns based on GO annotations related to brain development and neurogenesis. Only the human-specific differentially expressed genes with consistent human–chimpanzee or human–macaque differential expression detected in at least one of the three C1-based scRNA-seq datasets are shown. **l**, Ventral telencephalon cell heterogeneity in organoids was investigated by *t*-SNE embeddings with RSS profiles of human (3,385) and chimpanzee ventral (773) pseudocells combined as the input. Pseudocell clusters were annotated on the basis of marker gene expression. Pseudocells were also coloured by species and diffusion map based on medial ganglionic eminence (MGE) neuron developmental pseudotimes. **m**, *t*-SNE plots coloured by marker gene expression and in situ hybridization images from the Allen Developing Mouse Brain Atlas (available from https://developingmouse.brain-map.org/) showing expression of *Dlx5*, *Isl1* and *Sox6* in the mouse developing ventral forebrain embryonic day 13.5 (E13.5). **n**, Human–chimpanzee ventral differentially expressed genes are largely shared along the dorsal forebrain developmental trajectories. **o**, Human–chimpanzee DE directionalities and magnitudes and DE genes detection rates on the two trajectories. DE directionalities and magnitudes are consistent on the dorsal and MGE trajectories, with most of the shared DE genes showing the highest humanchimpanzee expression divergence at NPC. DE genes specifically detected on one trajectory have the tendency of higher detection rates on the trajectory where human–chimpanzee differential expression is detected.



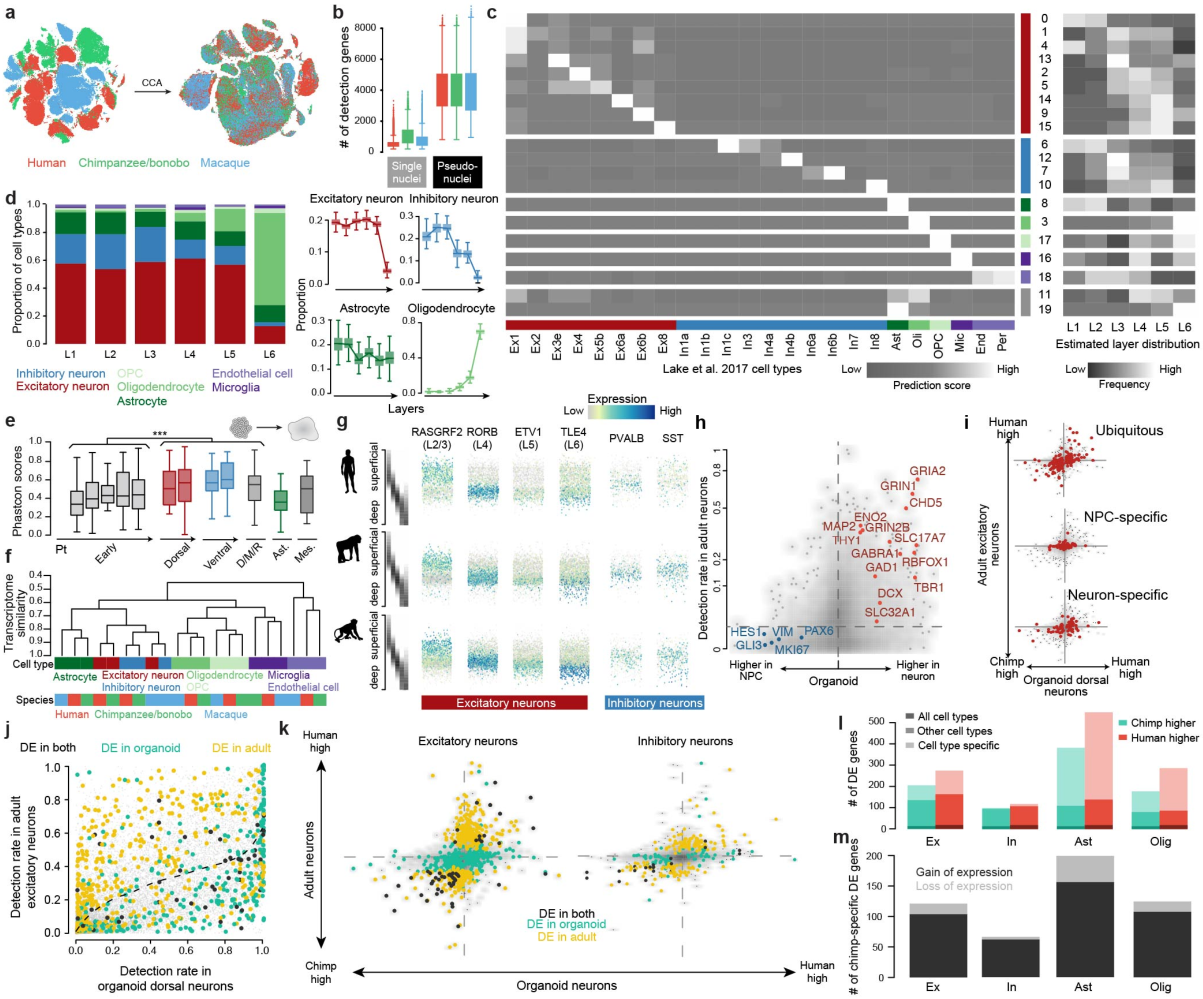
Extended Data Fig. 8: Chromatin accessibility in cerebral organoids during development. **a**, *t*-SNE projection of highly variable gene expression in Fluidigm C1-based scRNA-seq data of cerebral organoids.

Cortical cells are coloured red, with larger points corresponding to cells with paired expression and chromatin accessibility data (data generated from the same cell suspension). A total of 94.4% (219 out of 232 cells) of cells with paired data are cortical, validating the cortical origins of the dissected cerebral organoid regions. **b**, Cerebral organoid accessible peaks are significantly and highly enriched (Fisher's exact test) for overlapping human VISTA enhancers active in the forebrain relative to all other tissues (left). Three representative human VISTA enhancers with validated activity in E11.5 mouse forebrain that overlap cerebral organoid peaks (out of 268 such enhancers). **c**, Percentage of genes with accessible chromatin at the promoter of genes that are expressed or not expressed in human cerebral organoids. **d–f**, *t*-SNE projection of bias-corrected deviations in accessibility for 7-mers within organoid scATAC-seq peaks per cell, with cells colour coded by cell state (NPC, neuron) for human (**d**, 221 cells), chimpanzee (**e**, 543 cells), and macaque (**f**, 118 cells). Binding motif deviation Z-scores for representative transcription factors are shown, as well as deviation Z-scores for overlapping DA snATAC-seq peaks in mouse developing forebrain excitatory neurons [291]. **g**, Signal intensity tracks of aggregated and individual single-cell chromatin accessibility data per cell state in human organoids at a NPC-specific promoter peak (left) and a neuron-specific promoter peak (right). For comparison, cerebral organoid bulk ATAC-seq chromatin accessibility data and human fetal brain bulk DNase-seq is shown. **h**, Enrichment of representative enriched biological process GO terms associated with human NPC DA peaks (gold) or human neuron DA peaks (light red) relative to all human organoid accessible peaks. Each point in the scatter plot represents a GO term and is coloured by their enrichment in NPCs (yellow), neurons (red), both (dark red) or neither (grey). **i**, *t*-SNE plots coloured by pseudotime, and heat maps showing binding motif deviation Z-scores for chosen transcription factors (rows) in all cells (columns) ordered in pseudotime for human (left) and chimpanzee (right). **j**, *t*-SNE projection of bias-corrected deviations in accessibility for 7-mers within scATAC-seq peaks per cell (518 cells), with cells colour coded by time point, and organoid data colour coded by cell state. Binding motif deviation Z-scores for representative transcription factors are shown to the right. **k**, *t*-SNE plot with cells coloured by their deviation Z-score for overlapping differentially accessible snATACseq peaks from mouse developing forebrain [291] radial glia cells (left) or excitatory neurons (right). **l**, Diffusion map projection using the top-250 differentially accessible peaks per time point or cell state. The principle curve fit through the cells is shown as a black line. **m**, Proportion of cells scaled by row for each time point or cell state over pseudotime. **n**, Heat map representing the deviation Z-score of transcription factor motifs that significantly vary over the time course plotted for each cell across pseudotime.



Extended Data Fig. 9: Chromatin accessibility differences in human and chimpanzee cerebral organoids. **a**, Signal intensity tracks of aggregated single-cell and bulk chromatin accessibility data from human, chimpanzee and macaque at a human-specific NPC-specific DA peak (left) and a human-specific neuron-specific DA peak (right). **b**, The eight most significant human–chimp organoid DA peaks containing a fixed SNC and accessible only in the cerebral organoid stage that overlap a VISTA human enhancer with validated activity in the developing mouse forebrain (out of 68 such cases). For each DA peak, the accessibility across pseudotime is shown for human and chimpanzee with heat maps depicting cells where the peak is accessible (yellow) or inaccessible (black). The activity pattern of the overlapping VISTA enhancer in E11.5 mouse embryos is shown to the right. **c**, The proportion of DE genes (dark colour) or all expressed genes as background (light colour) with a human–chimp organoid DA peak overlapping the promoter region (blue) or is distal to the promoter region (pink). The plot shows that DE genes between human and chimpanzee organoids are more likely to have a nearby DA peak than background. **d**, Fixed SNCs predicted to significantly alter transcription factor binding within human–chimp organoid DA peaks, with the name of the altered motif shown for peaks linked to DE genes (red points). On the right, signal intensity tracks for a human motif gain (top) and human motif loss (bottom) within a human–chimp DA peak. **e**, Altered transcription factor motifs grouped by family plotted for their alteration rate, which is the number of times a family member's motif is altered in human–chimp organoid DA peaks divided by the number of times it is detected in all accessible organoid peaks. **f**, Twenty transcription factors with the highest alteration rate, which is the number of times a motif is altered in human–chimp organoid DA peaks divided by the number of times it is detected in all accessible organoids peaks.

Heat maps show their expression level in human and chimpanzee NPCs and neurons, with the bars to the left representing the average expression level across NPCs and neurons. **g**, Example of an accessible peak in chimpanzee and macaque that overlaps a computationally verified, non-polymorphic human conserved deletion (hCONDEL). **h**, **i**, Signal intensity tracks of aggregated single-cell or bulk chromatin accessibility data from human, chimpanzee and macaque for two genes, *LYPD1* (**h**) and *RAC1* (**i**), that have higher expression and specifically in humans, with genomic regions with gain of accessibility detected specifically in humans. Gene expression is shown in heat maps (bottom).



Extended Data Fig. 10: Supplementary analysis of human, chimpanzee and macaque adult brain snRNA-seq. **a**, The snRNA-seq data of adult brains in human (50,035), chimpanzee and bonobo (33,847) and macaque (50,403) were integrated using Seurat v.3. **b**, Box plots (boxes show IQR and whiskers show 1.5 x IQR) showing the number of detected genes in single nuclei and pseudonuclei (3,420 human, 3,831 chimpanzee and 4,623 macaque pseudonuclei). **c**, Heat map showing the average prediction scores of each of the 20 identified clusters to each of the cell types reported [215] by Lake et al. (2016), as well as their estimated distributions in different cortical layers in humans. Clusters are grouped in major cell classes. **d**, Cell type composition of layers and layer distribution of cell types in human. Left, stacked bars showing the estimated cell type composition of different layers. Right, box plots (boxes show IQR and whiskers show 1.5 x IQR) showing the estimated proportion per layer for four cell classes: excitatory neurons, inhibitory neurons, astrocytes and oligodendrocytes. **e**, Genomic conservation based on average phastCon scores of developmental stage markers (in total 2,000 genes) from iPSCs to neurons in human cerebral organoids (***$P < 10^{-10}$, two-sided Wilcoxon's rank-sum test, $n_1 = 818$ genes, $n_2 = 188$ genes). **f**, Hierarchical clustering of the average transcriptome of seven cell classes in the three species. **g**, Expression of layer markers (*RASGRF2*, *RORB*, *ETV1* and *TLE4*) in excitatory neurons and inhibitory neuron subtype markers (*PVALB* and *SST*) in inhibitory neurons, along the predicted laminar origin of the pseudonuclei in human, chimpanzee and bonobo, and macaque. **h**, Detection rate in adult tissue of genes being differentially expressed between NPCs and neurons in organoids. **i**, Comparison of human–chimpanzee DE in adult excitatory neurons and that in organoid dorsal neurons for the robust DE genes detected in the organoid dorsal forebrain trajectory. Three categories of DE genes are highlighted: ubiquitous DE in organoids (top), DE only in NPCs (middle) and DE only in neurons (bottom).

j, Comparison of gene detection rates in organoid dorsal neurons and adult excitatory neurons, with human–chimpanzee DE genes in adult excitatory neurons coloured in yellow, DE genes in organoid dorsal neurons coloured in green, and shared DE genes coloured in black. The dashed curve shows the fitted relationship between the two systems using all genes. Area below the curve represents higher detection rate in organoid neurons than adult neurons and area above the curve represents higher detection rate in adult neurons. **k**, Comparison of human–chimpanzee DE (left) between organoid dorsal neurons and adult excitatory neurons, and between organoid ventral MGE neurons and adult inhibitory neurons (right). Densities are shown as grey scale shadows, with human–chimpanzee DE genes highlighted (yellow, DE only in adult; green, DE only in organoids; black, DE in both). **l**, Number of human and chimp DE genes for cell classes based on all cell types, a subset of cell types and specific cell types. **m**, Number of chimpanzee-specific DE genes across cell classes. The majority of the chimpanzee-specific DE genes have gain of expression (dark) rather than loss of expression (light).

Supplementary information

SUPPLEMENTARY MATERIALS

SUPPLEMENTARY TABLES

Supplementary tables can be found in the online version of the article (doi: 10.1038/s41586-019-1654-9).

Supplementary Table 1: Overview of single-cell genomic experiments

Supplementary Table 2: Highly variable genes used in SPRING and Reference Similarity Spectrum analyses. Highly variable genes were determined using Seurat (v2.3, vst method)

Supplementary Table 3: Metadata annotations of human cells from scRNA-seq

Supplementary Table 4: Cluster marker genes from heterogeneity analysis of human whole trajectory data. Markers were identified using Seurat (v2.3, two-sided Wilcoxon's rank sum test, Bonferroni corrected $P<0.05$)

Supplementary Table 5: Metadata annotations of chimpanzee cells from scRNA-seq

Supplementary Table 6: Cluster marker genes from heterogeneity analysis of chimpanzee whole trajectory data. Markers were identified using Seurat (v2.3, two-sided Wilcoxon's rank sum test, Bonferroni corrected $P<0.05$)

Supplementary Table 7: Expression of human-specific genes in human whole trajectory data

Supplementary Table 8: Metadata annotations of differentially expressed genes between human, chimpanzee, and macaque in cerebral organoids

Supplementary Table 9: Metadata annotations of cells from scATAC-seq

Supplementary Table 10: Metadata annotations of differentially accessible peaks between human and chimpanzee cerebral organoids

Supplementary Table 11: Metadata annotations of peaks overlapping human conserved deletions and single nucleotide changes in differentially accessible peaks

Supplementary Table 12: Highly variable genes used in adult brain single-nuclei RNAseq data integration from human, chimpanzee and macaque. Highly variable genes were determined using Seurat (v3.0, vst method).

Supplementary Table 13: Metadata annotations of nuclei and pseudonuclei from snRNA-seq

Supplementary Table 14: Major cell class marker genes of human, chimpanzee and macaque adult brain. Markers were identified using Seurat (v3.0, two-sided Wilcoxon's rank sum test, Bonferroni corrected $P<0.05$)

Supplementary Table 15: Metadata annotations of differentially expressed genes between human, chimpanzee and macaque in adult brains

SUPPLEMENTARY NOTES

Progression of major cell types during cerebral organoid culture We used the time course single-cell RNA-seq data from cerebral organoids to track a progression through pluripotent, neuroectodermal, and neuroepithelial stem cell states during the first 15 days of differentiation. By 1 month, cells diversify into neural progenitors from multiple brain regions including the forebrain (dorsal and ventral telencephalon, diencephalon), midbrain (mesencephalon), and hindbrain (rhombencephalon). A small subpopulation resembling retinal progenitors of the developing eye field is also present, but these cells were only detected in an iPSC 409b2-derived organoid. In addition, a nonneuronal mesenchymal population appears from both cell lines early in the differentiation time course. By 2 months, excitatory and inhibitory neuronal fates have differentiated from progenitors of multiple brain regions, and by 4 months astrocytes have emerged.

Focusing on 2-month organoids, we observed neuronal differentiation trajectories representing ventral and dorsal telencephalon, as well as distinct populations of cortical excitatory (GLI3, EOMES, NEUROD6), ventral telencephalon inhibitory (DLX1, SOX6, GAD1/2), diencephalon excitatory, diencephalon inhibitory (with Cajal-Retzius cell signatures), mesen- (or midbrain) and rhombencephalon (hindbrain) excitatory, and mesenand rhombencephalon inhibitory neurons. Each iPSC line contributed cells to multiple differentiation trajectories, however the proportions of cells in each trajectory varied across organoid and iPSC. For example, over 90% of cells from the line Kucg2 were on the cortical excitatory (dorsal) trajectory in each of the 3 organoids, whereas Hoik1-derived organoids predominantly contained cells from non-telencephalic regions.

In the single-cell RNA-seq data we generated from human, chimpanzee, and macaque 2 and 4 month organoids using 10X Genomics, we observed very strong signatures of deep and upper layer cortical neuron differentiation in chimpanzee and macaque organoids. This bifurcation of deep and upper layer cortical neurons was much less pronounced in human organoids from the same time point. We analyzed additional scRNA-seq data (Smart-seq2, Fluidigm C1) data from 52 human organoids from 15 individuals, 38 chimpanzee/bonobo organoids from 11 lines, and one macaque organoid from one line. Based on this data, we find that there is variation in the degree of specification of upper and deep layer cortical neurons, which makes it unclear whether this is a reliable measure of organoid maturation. It is unclear if this variation is due to batch, lines, organoid or scRNA-seq protocols. However, we found that neuron projection, synapse, and neurotransmitter related genes were consistently and significantly expressed higher at an earlier time point in the chimpanzee organoids compared to the human organoids, suggesting that human organoids mature slower than those of other primates.

Potential biological significance of chosen genes with differential properties between human and chimpanzee We compared scATAC-seq data of human and chimpanzee cerebral organoids and identified 8,099 peaks (7.4% of all accessible peaks) that gained accessibility in humans relative to chimpanzee, whereas 9,836 peaks (9% of all accessible peaks) lost accessibility. We annotated peaks that are DA between humans and chimpanzees with various evolutionary signature. For instance, we identified 62 human accelerated regions that overlap DA peaks (32 gain accessibility in human, 30 gain accessibility in chimpanzee), with one of these sites being nearby a gene with human-specific expression. In this case, the potential regulatory region is 244 Kb away from cadherin 7 (CDH7), a gene with higher expression specifically in human cortical neurons, and has increased accessibility in human neurons relative to chimpanzee and macaque. We also find DA regions nearby two genes, Ly6/PLAUR domain-containing protein 1 (LYPD1) and Ras-related C3 botulinum toxin substrate 1 (RAC1), that have human-specific expression in NPCs and neurons, respectively. LYPD1 is involved in neurotransmitter receptor-binding and anxiety-related behaviors [307] and RAC1 is a GTPase involved in diverse processes including glucose uptake and cytoskeletal reorganization and genetic variants in this gene can lead to micro-or macrocephaly [308]. In addition, we identify 22 regions that are accessible in chimpanzee NPCs or neurons that are highly conserved in mammals, but the DNA has been deleted in humans (so-called human conserved deletion, hCONDELs) [60] and 1 of these are located nearby a DE gene (FADS1). FADS1 encodes the fatty acyl desaturases (delta-5 desaturase) which catalyze key steps in the ω-3 and ω-6 lipid biosynthesis pathways [309], and has been reported to be abundantly expressed in the brain [310], and therefore could have contributed to the fast divergence of lipids between human and other primates like chimpanzee and macaque [52].

With the single-nucleus RNA-seq data of human, chimpanzee and macaque adult prefrontal cortex, we identified 479 DE genes in adult excitatory neurons between human and chimpanzee. Among them, 53 overlap with DE genes in dorsal telencephalon cells between human and chimpanzee organoid. They include COL6A1, the gene encoding alpha1-chain of collagen VI, the broadly distributed extracellular matrix protein. COL6A1 has been suggested to play roles in central nervous system development and diseases [279], and has been shown to have a protective role limiting autophagy and apoptosis in aging neurons [311]. Another example is RIC3, which shows not only human-specific DE in adult neurons, but also human-specific DE in organoids in a neuron-specific manner. RIC3 encodes a protein which functions as a chaperone influencing the folding, assembly of specific 5-hydroxytryptamine type 3 receptor and nicotinic acetylcholine receptor subtypes, and regulates the number and maturation of acetylcholine-gated ion channels in neurons [312].

Chapter 3

Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia

Johannes Klaus, Sabina Kanton, Christina Kyrousi, Ane Cristina Ayo-Martin, Rossella Di Giaimo, Stephan Riesenberg, Adam C. O'Neill, J.Gray Camp, Chiara Tocco, Malgorzata Santel, Ejona Rusha, Micha Drukker, Mariana Schroeder, Magdalena Götz, Stephen P. Robertson, Barbara Treutlein and Silvia Cappello

Nature Medicine, Vol. 25, Issue 4 (2019)

CHAPTER 3

Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia

Johannes Klaus[1,11], Sabina Kanton[2,11], Christina Kyrousi[1,11], Ane Cristina Ayo-Martin[1,3], Rossella Di Giaimo[1,4], Stephan Riesenberg[2], Adam C. O'Neill[5,6], J. Gray Camp[2], Chiara Tocco[1], Malgorzata Santel[2], Ejona Rusha[7], Micha Drukker[7], Mariana Schroeder[1], Magdalena Götz[6,8], Stephen P. Robertson[5], Barbara Treutlein[2,9,10]* and Silvia Cappello[1]*

Affiliations:

[1]Max Planck Institute of Psychiatry, Munich, Germany.

[2]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

[3]International Max Planck Research School for Translational Psychiatry (IMPRS-TP), Munich, Germany.

[4]Department of Biology, University of Naples Federico II, Naples, Italy.

[5]Department of Women's and Children's Health, University of Otago, Dunedin, New Zealand.

[6]Institute of Stem Cell Research, Helmholtz Center Munich, Munich, Germany.

[7]Institute of Stem Cell Research, iPSC Core Facility, Helmholtz Center Munich, Munich, Germany.

[8]Ludwig Maximilian University, Munich, Germany.

[9]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany.

[10]Technical University Munich, Department of Biosciences, Freising, Germany.

[11]These authors contributed equally: Johannes Klaus, Sabina Kanton, Christina Kyrousi.

Correspondence to: barbara_treutlein@eva.mpg.de,
silvia_cappello@psych.mpg.de

---

[11] These authors contributed equally to this work

Abstract

Malformations of the human cortex represent a major cause of disability [79]. Mouse models with mutations in known causal genes only partially recapitulate the phenotypes and are therefore not unlimitedly suited for understanding the molecular and cellular mechanisms responsible for these conditions [313]. Here we study periventricular heterotopia (PH) by analyzing cerebral organoids derived from induced pluripotent stem cells (iPSCs) of patients with mutations in the cadherin receptor–ligand pair *DCHS1* and *FAT4* or from isogenic knockout (KO) lines [79, 227]. Our results show that human cerebral organoids reproduce the cortical heterotopia associated with PH. Mutations in *DCHS1* and *FAT4* or knockdown of their expression causes changes in the morphology of neural progenitor cells and result in defective neuronal migration dynamics only in a subset of neurons. Single-cell RNA-sequencing (scRNA-seq) data reveal a subpopulation of mutant neurons with dysregulated genes involved in axon guidance, neuronal migration and patterning. We suggest that defective neural progenitor cell (NPC) morphology and an altered navigation system in a subset of neurons underlie this form of PH.

Main text

Mammalian neocortical development represents a highly orchestrated process that depends on the precise generation, migration and maturation of neurons. The importance of a coordinated sequence is underlined by the conditions with its disruption: malformation of cortical development. PH represents one of the most common forms of these disorders and is characterized by heterotopic neurons lining their sites of production. Patients with PH typically present with intellectual disability, and this is frequently associated with epilepsy [226, 314–319]. The identification of mutations in the protocadherins DCHS1 and FAT4 put the spotlight on defects in NPCs as a causal mechanism of the condition. Here we explore the functions of DCHS1 and FAT4 in the developing cortex using human iPSC-derived NPCs, neurons and cerebral organoids.

We first reprogrammed fibroblasts from control individuals and patients with PH who carry mutations in *DCHS1* or *FAT4* into iPSCs (Extended Data 1a,b). Specifically, fibroblasts were collected from two different previously characterized patients [79, 227]: one was compound heterozygous for mutations in the *FAT4* gene and one homozygous for mutation in the *DCHS1* gene. Additionally, to control for differences due to the different genomic background in the patients, we generated KO iPSC lines for both genes using CRISPR–Cas9 genome editing in control iPSCs. We programmed the iPSCs toward NPCs and neurons in two-dimensional (2D) culture and generated three-dimensional (3D) cerebral organoids [250] (Extended Data 1c,d). Using in situ hybridization, we identified that both genes were expressed in the periventricular structures of cerebral organoids and neurons (Extended Data 1e–h'), a pattern consistent with that detected in mouse and human [79, 320]. These findings were confirmed by scRNA-seq of cells derived from cerebral organoids, where the expression of *DCHS1* and *FAT4* was found in both progenitors and neurons (Extended Data 1i,j).

To investigate whether PH is recapitulated within cerebral organoids (Extended Data 1k,l), we scrutinized ventricular zone structures for such phenotypes. In control organoids, a clear distinction of the neuronal (MAP2+ cells or NEUN+ cells) layer from the germinal zone (PAX6+ cells) was identified (Fig.1a,d, Extended Data 2a,e,h and Extended Data 3g,k). Organoids derived from mutant or KO iPSC lines exhibited a significant number of neuronal nodules at ventricular positions (Fig.1a–c, Extended Data 2e–j, and Extended Data 2l–l"'). In addition to this neuronal heterotopia-like phenotype, mutant and KO organoids presented poorly organized germinal zones (Fig.1d,e and Extended Data 2d). This feature was especially apparent in germinal zones of *FAT4*-mutant organoids, with most not showing evident separation of the neuronal band from the germinal zone, with neurons intruding in most cases. In the case of *DCHS1*-mutant or knockdown organoids, performed via electroporation of specifically designed microRNAs (miRNAs) that target the human gene (Extended Data 2k), clusters of neurons were found within the germinal zones where NPC processes where disrupted (Fig.1f–f"' and Extended Data 2l–l"'). Although *DCHS1*-mutant organoids displayed clearer separations between the germinal zones and neuronal layer, the neurites showed an altered morphology compared with controls, with many processes appearing as thick bundles (Extended Data 2m–o). These findings show that cerebral organoids recapitulate the hallmarks of PH. The loss of organization (germinal zone versus cortical plate) in mutant organoids starts early, after 20 d of cerebral organoid development (Extended Data 3a–f), even though the sizes of the two zones (VZ and CP) are not altered (Extended Data 3g–j).

In mice, knockdown of *Dchs1* or *Fat4* leads to overproliferation of progenitors1. This increase in proliferation was not observed in human organoids derived from patient iPSCs (Extended Data 3k–s), highlighting fundamental species-specific differences. We analyzed the morphology of the processes of NPCs in organoids by NESTIN immunostaining in mutant or KO organoids. Whereas in control organoids the processes appeared to be aligned and straight, in *FAT4*-mutant or KO organoids, these processes were often disrupted and exhibited a twisted morphology (Fig.1g,i,m and Extended Data 4a,c). NPC morphology in *DCHS1*-mutant or KO organoids was less severely compromised (Fig.1g–i,m and Extended Data 4a–c). To scrutinize the morphology of single NPCs, we elec- troporated a Gap43-GFP plasmid [321] to label the cell membrane in mutant organoids. Analysis 4 d after electroporation showed that *FAT4*-mutant organoids had disrupted progenitor morphology (Fig.1j,l,n). Consistent with previous results, this phenotype was also observed in *DCHS1*-mutant organoids (Fig.1j,k,n).

To ensure that the observed phenotypic differences are not caused by the different genetic backgrounds of the cells of origin, we respectively and selectively knocked down *DCHS1* and

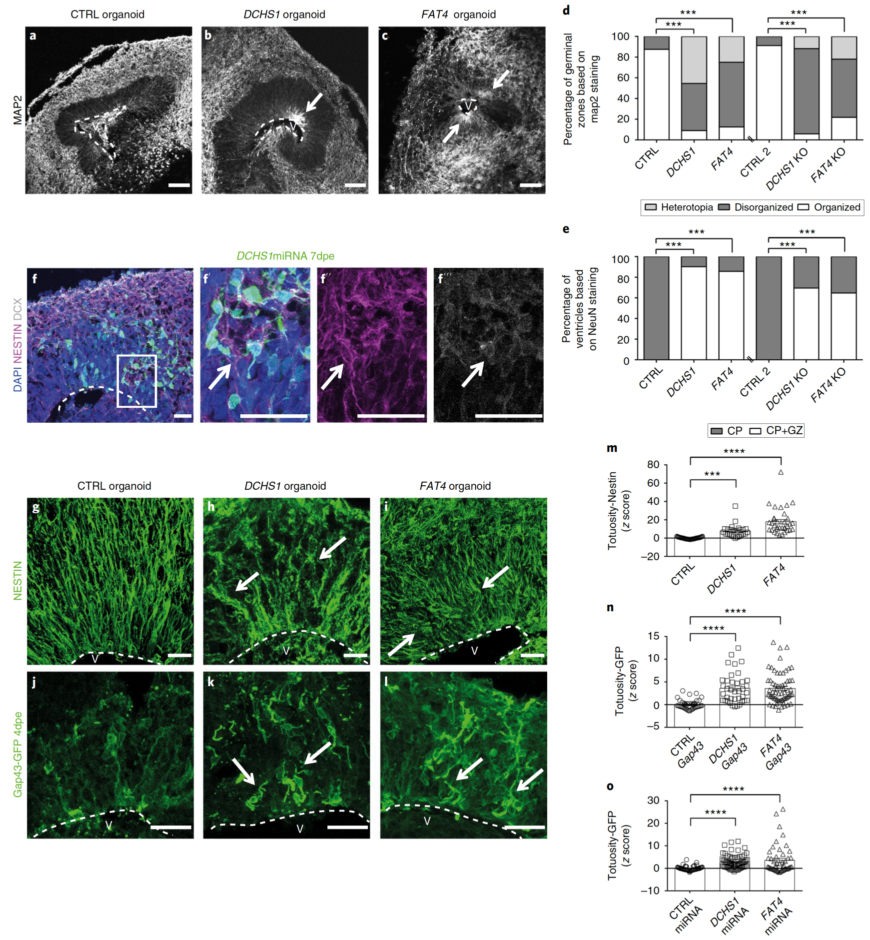


Fig. 1: Mutations in DCHS1 and FAT4 cause neuronal heterotopia and disturbances in the morphology of NPCs in cerebral organoids. **a**–**c**,**f**–**f‴**, Micrographs of sections of organoids immunostained as indicated in the panels. Note the mispositioning of neurons marked by arrows in mutant or in electroporated cerebral organoids. **a**–**c**, batches (b) $= 5$, organoids (o) $= 15$ per condition, **f**–**f‴**, $b = 2$, $o = 6$ per condition. **d**,**e**, Classification of organoids according to phenotypes ($v = 24$ CTRL, 11 *DCHS1*, 16 *FAT4*, $v = 23$ CTRL2, 17 *DCHS1* KO, 32 *FAT4* KO based on MAP2 staining; $v = 7$ CTRL, 31 *DCHS1*, 14 *FAT4*, 12 CTRL2, 23 *DCHS1* KO, 17 *FAT4* KO based on NEUN staining). Mutant or KO organoids show significantly more disorganized germinal zones and exclusively show a heterotopia phenotype (chi-squared test, $\chi^2{}_{(4)} = 84.79$, $P = 0.000$ for **d** in mutant organoids, chi-squared test, $\chi^2{}_{(4)} = 40.33$, $P = 0.000$ for **d** in KO organoids, chi-squared test,$\chi^2{}_{(4)} = 27.69$, $P = 0.000$ for **e** in mutant organoids, Chi square test, $\chi^2{}_{(4)} = 16.94$, $p = 0.000$ for **e** in KO organoids). **b**,**c**, Ectopic nodule of MAP2+ neurons in the germinal zone of *DCHS1* and *FAT4*-mutant organoids. **f**–**f‴**, Nodule of DCX+ neurons intermingling with NESTIN+ processes of NPCs in the germinal zone after *DCHS1* miRNA electroporation in control organoid. Dotted lines highlight ventricles (V). **g**–**i**, Micrographs of sections of mutant cerebral organoids (day 42) immunostained for NESTIN. Arrows indicate the disrupted morphology of NPCs. **g**–**l**, $b = 3$, $o = 9$ per condition. **j**–**l**, Micrographs of sections of mutant organoids electroporated with Gap43-GFP at day 42 and analyzed at day 46. Progenitors in mutant organoids (*k*,*l*) show disrupted pattern (arrows). **m**–**o**, Quantification of the tortuosity index of NPCs based on NESTIN ($v = 8$ CTRL ($c = 43$), 4 *DCHS1* ($c = 38$), 6 FAT4 ($c = 65$)) (m) or GFP ($v = 9$ CTRL ($c = 44$), 5 *DCHS1* ($c = 25$), 7 *FAT4* ($c = 34$)) (**n**) staining in mutant organoids, or based on GFP staining in control organoids electroporated with miRNAs against *DCHS1* or *FAT4* ($v = 11$ CTRL ($c = 52$), 13 *DCHS1* ($c = 71$), 7 *FAT4* ($c = 49$)) (o). Results are z scores; significance is based on one-way ANOVA; $P = 0.000$; Holm–Sidak multiple-comparisons test was performed for defining statistical differences between the three genotypes in each graph. Data in graphs are represented as mean ± s.e.m. Scale bars, 100 µm in **a**–**c** and 30 µm in **f**–**f‴** and **g**–**l**. Dotted lines highlight ventricles (V). $^{***}P < 0.001$, $^{****}P < 0.0001$.

FAT4 in NPCs via electroporation of miRNAs in control organoids. 7 d post electroporation, we similarly observed a disruption in the morphology of NPCs upon *DCHS1* and *FAT4* knockdown (Fig. 1o and Extended Data 4d–f), suggesting that the morphology of NPCs is disturbed in organoids derived from cells with mutations or reduced *DCHS1* and *FAT4*.

Dchs1 and Fat4 have been described to localize apically and to form a heterophilic complex in the developing mouse brain [320]. We therefore hypothesized that DCHS1 and/or FAT4 could be key molecules in the maintenance of the polarity of NPCs. We found mislocalization of cilia labeled by ARL13B (Extended Data 5r), but the total proportion of apical and basal cilia was not altered (Extended Data 5p–s); in FAT4-mutant germinal zones and, to some extent, in DCHS1-mutant germinal zones, the apical surface exhibited a disrupted or discontinuous morphology with increased distance between the apical surface and the NPCs (Extended Data 5f–o). This finding further underlined that the defects in NPCs are more pronounced in *FAT4*-mutant organoids, but it also revealed that both mutations do not preclude the establishment of apicobasal polarity. To evaluate whether cytoskeletal dynamics are affected and possibly contributing to the disrupted morphology of NPCs in the patient-derived organoids, we performed immunohistochemistry for ACETYLATED TUBULIN, which labels stable microtubules. While the stable microtubules were not strongly affected in *DCHS1*-mutant NPCs, *FAT4*-mutant NPCs showed not only decreased organization but also markedly diminished levels of stable ACETYLATED TUBULIN (Extended Data 5a–c). To specifically detect levels of stable microtubules in NPCs, we performed a western blot analysis of ACETYLATED TUBULIN in NPCs from monolayer cultures and detected a significant decrease in ACETYLATED TUBULIN in *FAT4*-mutant NPCs. In contrast, TYROSINATED TUBULIN, which reflects the more dynamic form of tubulin, was not altered (Extended Data 5d,e).

In order to gain deeper insight into the molecular changes in progenitors and neurons of *DCHS1* and *FAT4* organoids, we next performed scRNA-seq to dissect the cell composition and transcriptional landscapes in *DCHS1*, *FAT4* and control (CTRL) organoids. We analyzed 805 single cells from three microdissected organoid cortical regions for each condition (nine organoids in total; 50–60 d) and identified progenitors and neurons based on known marker genes (Fig. 2a,b). In an unbiased principal component analysis (PCA), the first two components separate NPCs (e.g. *VIM*, *PAX6*, *SOX2*, *HES1*) and neurons (e.g. *DCX*, *STMN2*, *TUBB3*, *SNAP25*, *MYT1L*) independent of the condition (control, *DCHS1*, *FAT4*) (Fig. 2a,b and Supplementary Table 1). However, the third principal component segregates control cells from *DCHS1*- and *FAT4*-mutant cells, which remarkably intermingle, suggesting that cells derived from patients with mutation in *DCHS1* and *FAT4* share common transcriptome features distinct from the control cells (Fig. 2c). Interestingly, among the genes upregulated in all mutant cells, there are genes involved in patterning (*HOX* genes, *MEIS1*) as well as *RND3*, a negative regulator of *RHOA* that is essential for correct neurogenesis [228].

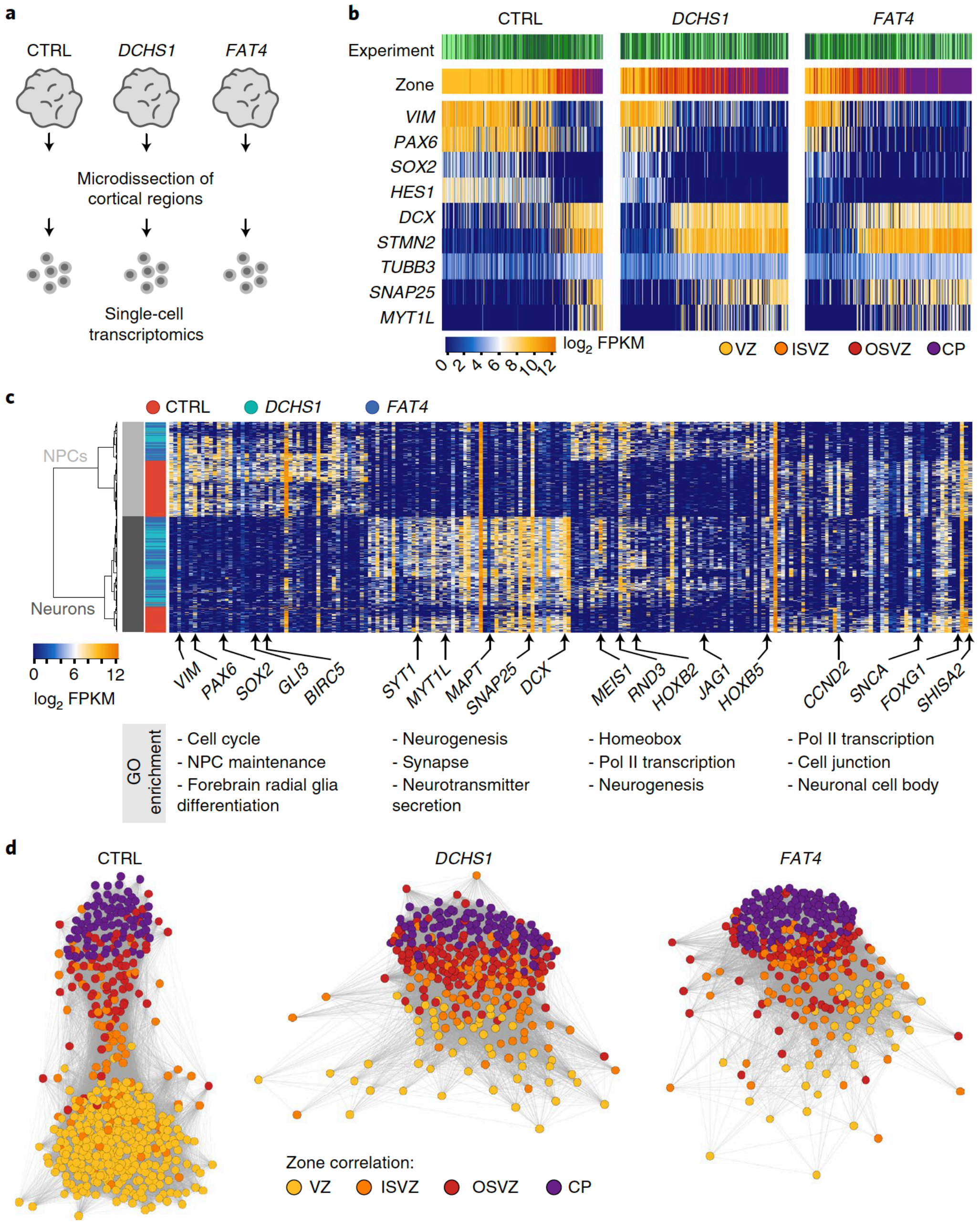


Fig. 2: scRNA-seq reveals alterations in mutant NPCs. **a**, Schematic of scRNA-seq experimental procedure. Number of organoids and cells analyzed: 9 organoids (3 for each condition), 805 single cells (316 CTRL, 255 *DCHS1*, 234 *FAT4*). **b**, Heat map showing expression of selected marker genes (rows) for NPCs and neurons for all three conditions (CTRL, *DCHS1*, *FAT4*), with single cells represented in columns. Top side bar visualizes experiment (shades of green) and maximal zone correlation for each single cell (VZ, ventricular zone, yellow; iSVZ, inner subventricular zone, orange; oSVZ, outer SVZ, red; CP, cortical plate, purple). Cells are ordered based on their PC 2 loading, corresponding to the trajectory from NPCs to neurons. **c**, Hierarchical clustering on all 805 single cells (rows) and genes (columns) identified by PCA to correlate and anticorrelate with PC 2 and PC 3 (Supplementary Table 1). The first two components commonly separate progenitors (light gray) from neurons (dark gray) for all experiments. The third component separates CTRL cells (red) from *DCHS1* (cyan) and *FAT4* (blue) cells. Results of gene ontology (GO) enrichment analysis for each gene group are shown below the heat map. **d**, Pairwise correlation network reveals a differentiation topology from NPCs correlating with the VZ (yellow) via cells correlating with iSVZ (orange) and oSVZ (red) to neuronal cells in the CP (purple) for CTRL cells (left), *DCHS1* cells (center) and *FAT4* cells (right).

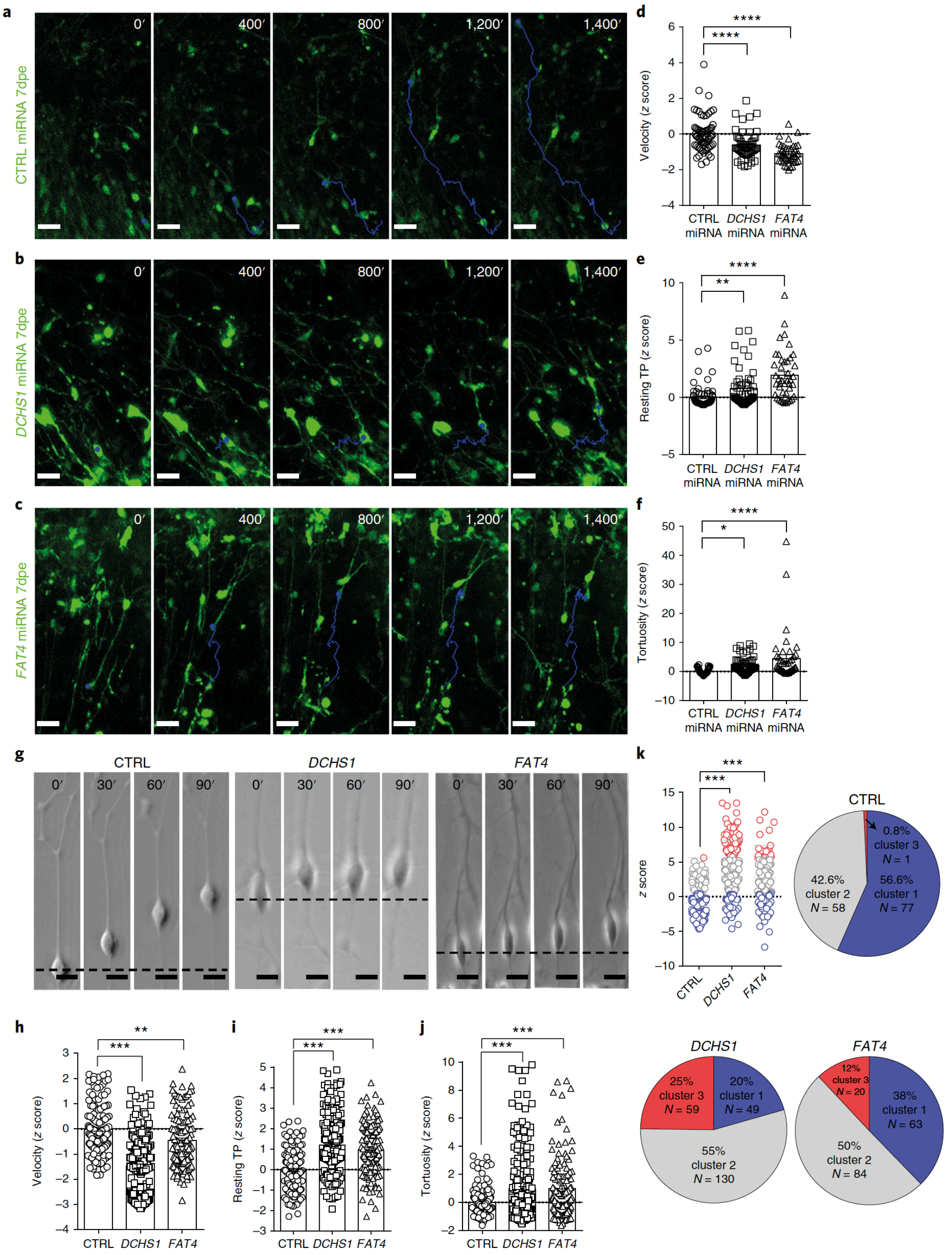


Fig. 3: Time-lapse imaging of mutant or knockdown neurons reveals an altered migration pattern. **a**–**c**,**g**, Examples of the movement of neuronal somata in control organoids after control, *DCHS1* or *FAT4* miRNA electroporation monitored in 3D time-lapse experiments (**a**–**c**) and in control and mutant neurons monitored in 2D time-lapse experiments (g). Scale bars, 30 μm in **a**–**c** and 10 μm in **g**. **a**–**c**, $b = 1$, $o = 3$ per condition; g, $b = 4$. **d**–**f**,**h**–**j**, Quantification of velocity, number of resting time points and tortuosity differed according to the mutation. **d**–**f**,**h**–**k**, CTRL miRNA $b = 1$, $o = 3$, $v = 7$, $c = 63$; *DCHS1* miRNA $b = 1$, $o = 3$, $v = 7$, $c = 56$; *FAT4* miRNA $b = 1$, $o = 3$, $v = 7$, $c = 42$; CTRL $b = 4$ $c = 136$; *DCHS1* $b = 4$ $c = 220$; *FAT4* $b = 4$ $c = 126$. Results are z scores; statistical analysis was performed in **d**,**e** based on one-way ANOVA, $P = 0.000$, Holm–Sidak's multiple comparisons test was performed for defining statistical differences between the three genotypes in each graph, in **h**–**j** based on one-way ANOVA with Tukey HSD post hoc and through multivariate ANOVA (**h**, $F(6{,}956) = 84.20$, $P = 0.000$; i, $F(6{,}956) = 43.10$, $P = 0.000$, **j**, $F(6{,}956) = 10.68$, $P = 0.000$). **k**, Hierarchical and two-step cluster analysis of migration dynamics across the whole cell population (chi-squared test, $\chi^2_{(2)} = 70.32$, $P = 0.000$). Data in graphs are represented as mean ± s.e.m. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

Previously, we identified a group of genes (Supplementary Table 2) that described the different cell populations in the human fetal neocortex and found that cerebral organoids recapitulate the expression pattern of these genes [114]. Using these genes, we reconstructed the lineage relationships between cells and visualized the relationships as intercellular correlation networks for control, *DCHS1* and *FAT4* organoids. As expected, control cells form a network with a topology that recapitulates the layered structure of the healthy developing cortex (Fig. 2d). In contrast, the networks from mutant organoids have a generally deformed architecture, consistent with the observed heterotopia and disorganization of mutant organoids (Fig. 1). Mutant progenitors from *DCHS1* and *FAT4* organoids showed a higher correlation with signatures of the cells located in the inner and outer subventricular zone, at the expense of ventricular zone signatures (Fig. 2b,d and Extended Data 3o). This finding supports the observed morphological changes in mutant progenitors, which prematurely delaminate, typical of more-differentiated basally located progenitors (Fig. 1f–o, Extended Data 2l–l"' and Extended Data 4). Consistent with these findings, more differentiated neurons were found in mutant organoids compared with control organoids in each experiment (Fig. 2b and Extended Data 3p). We further validated these results by analyzing the proportion of proliferating and differentiated cells in patient-derived cerebral organoids by means of FACS analysis (KI67, cycling progenitors; DCX, newborn neurons) from whole organoids (Extended Data 3q–s).

We also aimed at specifically characterizing the migration abilities of neurons with defective *DCHS1* or *FAT4*. To this end, we electroporated control organoids with specific miRNAs targeting *DCHS1* or *FAT4*. 7 d later, we tracked the migratory behavior of the electroporated neurons via time-lapse imaging in 300 µm-thick slices of the organoids with preserved 3D structure (Fig.3a–c). We measured speed of migration (velocity), time neurons spend without moving (resting time points) and ability to move in a straight direction (tortuosity), a typical feature of radially migrating neurons generated in the cerebral cortex. For all three parameters measured, the cells with *DCHS1* or *FAT4* knockdown displayed a significant difference compared with control cells. Specifically, the number of resting time points and the tortuosity were increased, whereas the velocity of the cells decreased (Fig. 3d–f). To investigate if this behavior is an intrinsic feature of the defective neu- rons or due to the changed NPC scaffold, which we observed in mutant and knockdown *DCHS1* and *FAT4* organoids (Fig. 1 and Extended Data 2,4), we performed time-lapse imaging in cultured mutant cells in a monolayer [322] (Extended Data 6a–f) and tracked the same parameters as those measured in 3D. In accordance with the behavior of migrating neurons observed in 3D, mutant neurons in 2D showed the same significant differences in their migration compared with controls (Fig. 3g–j). To see whether these changes affected only a specific population of cells, we performed hierarchical clustering analysis on the migration behavior of each cell based on all three parameters combined (velocity, resting time points and tortuosity). These data revealed that control cells cluster in two main distinct migration behaviors (cluster 1 and cluster 2). Remarkably, in mutant cells, cluster 2 remained the same as that in the control cells, whereas another type of migration behavior, cluster 3, emerged at the expense of cluster 1 (Fig. 3k). This change in the migratory dynamics of a specific cell population of patient-derived neurons could elucidate the phenotype observed in patients where only a limited number of neurons fail to migrate to their final location, resulting in heterotopic neurons residing at the ventricular surface.

We searched for a molecular signature that might explain the altered migration dynamics in a subset of mutant neurons. We used PCA and lineage reconstruction analysis (Monocle2) [323] of the organoid scRNA-seq data and found a subpopulation of neurons in *DCHS1*- and *FAT4*-mutant organoids that was not present in control organoids (Fig. 4a and Supplementary Table 3). This altered neuronal state consisting of mutant cells bifurcates from a progenitor-to-neuron trajectory that includes cells from all three conditions (Fig. 4b). Notably, this lineage bifurcation is not due to the emergence of neuronal cells with a different regional identity in the mutant organoids, because mutant, KO and control organoids express basic neuronal and cortical markers in a similar way (Extended Data 7a–m). The genes upregulated in the altered neuronal population are involved in fundamental neurodevelopmental processes like axon guidance, neuronal migration and patterning (e.g., *ROBO3*, *NDNF*, *DCC*, *CNTN2*, *EPHB3*, *EFNA3*, *ITGB1*, *GNG5*, *HOX*) (Fig. 4c). Interestingly, some genes such as GNG5 are expressed in progenitors, and are turned off in neurons from control organoids, but stay on in this altered neuronal state. Other genes such as *ROBO3* or *CNTN2* are only found in this subpopulation. On the contrary, the genes downregulated in these cells are involved in synapse formation, ion channel, axon guidance and cytoskeleton (e.g., *GRIA2*, *SNCA*, *MAPT*, *MAP1B*, *GNAI1*, *SPTAN1*, *FLRT2*). Interestingly, three of the dysregulated genes-*MAPT*, *MAP1B* and *SPTAN1*-are associated with epilepsy, a phenotype found in these patients [318, 324]. Variants in the *MAP1B* locus have also been associated with PH [318]. To further validate these data, we performed scRNA-seq (10X genomics) on neuronal cells derived from control or *DCHS1*-mutant cells in 2D cultures (Fig. 4d). Consistent with previous results, Monocle2 lineage reconstruction analysis revealed an altered neuronal state enriched for *DCHS1*-mutant cells (83% of altered neuronal cells) compared with control cells (17% of altered neuronal cells), which was characterized by a dysregulation of the same set of genes, as found in the organoid data (Fig. 4c), such as *GNG5*, *CNTN2* and *GRIA2* (Fig. 4e).Taken together, these findings suggest that the subpopulation of the heterotopically located neurons could be the consequence of alterations in migration dynamics due to incorrect regulation of specification and guidance gene programs. Similarly, not all neurons are affected in human patients with PH, but only a subset of them form periventricular clusters. It remains to be explored why the majority of neurons are able to navigate their way up to the cortical plate, despite the fact that they harbor the same gene mutation as the neurons located in the heterotopic clusters.

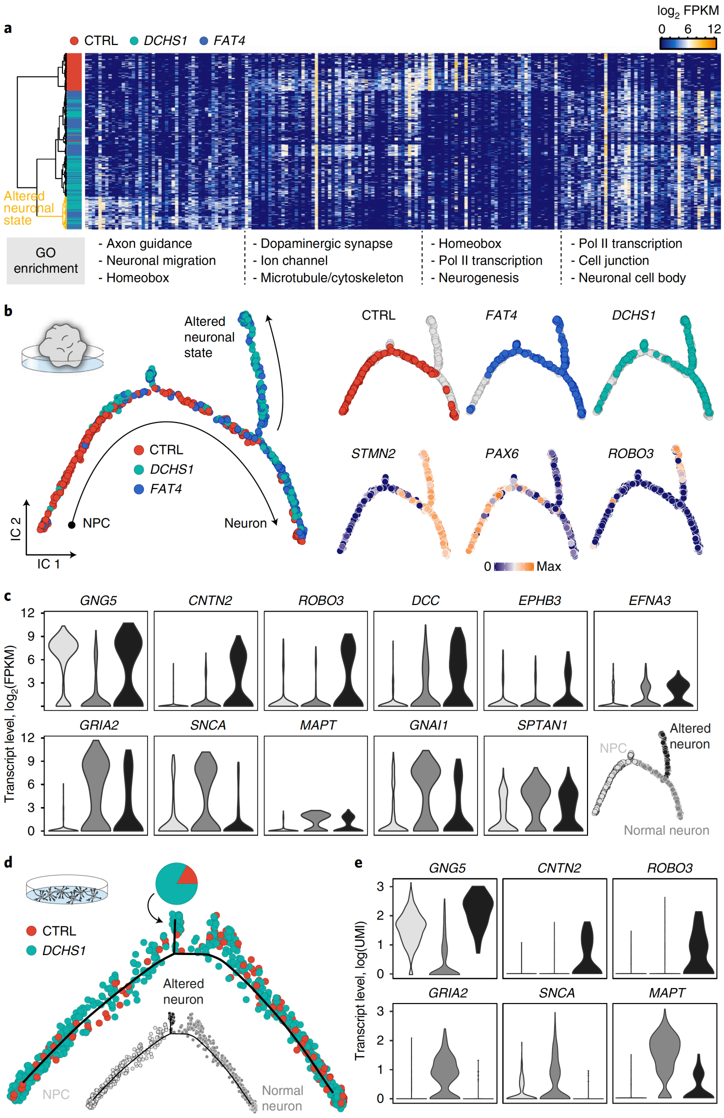


Fig. 4: scRNA-seq reveals an altered population of neurons. **a**, Hierarchical clustering on all neurons (rows, 467 single cells) and genes (columns) identified by PCA (Supplementary Table 3). A specific subpopulation of neurons (yellow) is identified that is only found in *DCHS1* or *FAT4* patient-derived organoids. Results of GO enrichment analysis for each gene group are shown below the heat map. **b**, Monocle2 lineage reconstruction of all 805 single organoid cells (circles) using genes identified by PCA (Supplementary Table 3) on all cells and on all neurons reveals a specific subpopulation of neurons only existing in mutant organoids that bifurcates from the general NPC-to-neuron trajectory. The three top right visualizations show cells from each condition separately in their respective color, with all other cells in gray. The three bottom right visualizations show the lineage reconstruction color coded by the expression of three selected genes marking neurons (*STMN2*), NPCs (*PAX6*) and the altered neuronal state (*ROBO3*). **c**, Violin plots show distribution of expression of genes that are up- (top) or downregulated (bottom) in the population of altered (black) compared with normal (dark gray) neurons. NPCs, light gray. Violin plots were plotted using ggplot2 default kernel density settings. **d**, Monocle2 lineage reconstruction of 10X genomics-based scRNA-seq data (437 CTRL, 751 *DCHS1*) from control and *DCHS1*-mutant neuronal cells grown in 2D culture reveals an altered neuronal subpopulation that bifurcates from the general NPC-to-neuron trajectory and is strongly enriched for mutant cells (83% of cells on this branch are *DCHS1* mutant). For the lineage reconstruction, the same genes were used as in **b**. **e**, Violin plots show distribution of expression of genes that are up- (top) or downregulated (bottom) in the population of altered (black) compared to normal (dark gray) neurons. NPCs, light gray. Note that these genes are the same as those found differentially expressed in the organoid altered neuronal state (**c**). IC, independent component.

Molecular and cellular analysis of organoids and cells derived from patients with mutations in *DCHS1* and *FAT4* or upon knock-down of these genes revealed a speciesspecific role of *DCHS1* and *FAT4* in human cells. Neuronal heterotopia develops as a consequence of multiple defects: the defective radial progenitor cells, which should guide the neurons to the correct final destination, and a specific subpopulation of neurons that acquire an altered navigation system, which changes their migratory dynamics and leads to compromised equipment for synaptic signaling. It is also clear that in patients, PH is induced by mutations in different genes, and thus it may involve the dysregulation of different molecular pathways [318].

Although the precise molecular regulation of *DCHS1* and *FAT4* remains to be clarified, we propose that changes in the expression of genes involved in the cytoskeleton and cell junction or axon guidance could affect the morphology of progenitor cells and migrating neurons, respectively.

The knowledge of the new molecular signatures acquired from disease neurons could ultimately be used to develop new strategies for targeting neurons that are aberrantly integrated into cortical circuits and possibly cause neuronal heterotopia in patients.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41591-019-0371-0.

Acknowledgements

We thank the families participating in this study for their involvement. We thank Y. Lu for help generating the microRNAs, M. Karow and I. Buchsbaum for helping with experiments and fruitful discussions in the lab, T. Öztürk for excellent technical support, A. Weigert for organoid culture, J. Kageyama for helping with data processing, R. Snabel for helping with Smart-seq2 libraries, the Core Unit Flow Cytometry at the Zentrum für Infektionsmedizin (veterinary faculty of the University of Leipzig) and the Core Unit Qualitätsmanagement/Technologieplattform at the Sächsischer Inkubator für Klinische Translation (SIKT) in Leipzig for karyotyping. This work was supported by funding from the DFG CA1205/2-1 (S.C.), ForIPS (M.G.), by the Max Planck Society (S.C., B.T.), by the Boehringer Ingelheim Fonds (S.K.), by the Health Research Council of NZ and Curekids (S.P.R.) and by an ERC Starting Grant (B.T.).

Author contributions

S.C. conceived and designed the research project, J.K., S.K., C.K., A.C.A.-M., R.D.G., S.R., A.C.O., C.T., M. Santel. and E.R. performed experiments and collected data, J.K., C.K., S.K., J.G.C., M. Schroeder, B.T. and S.C. analyzed data, M.D. reprogrammed patients' samples, M.G. was involved in the start of the project, contributed to data discussion and supervision of J.K. S.P.R. was involved in patient sample collection and critical discussion, J.K., S.K., C.K., B.T. and S.C. wrote the manuscript. All authors provided ongoing critical review of results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41591-019-0371-0.
Supplementary information is available for this paper at https://doi.org/10.1038/ s41591-019-0371-0.
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to B.T. or S.C.

Methods

Genetic mutation in DCHS1 and FAT4. Details about the genetic mutations found in *DCHS1* and *FAT4* patients can be found in the Supplementary Table 4.

iPSC culture. iPSCs were regularly maintained on Geltrex-coated dishes (Thermo Fisher, Waltham, MA, USA) in mTesR1 medium (Stem Cell Technologies, Vancouver, Canada) at 37°C, 5% CO_2 and ambient oxygen level. Passaging was done manually by scraping the cells after 10 min of collagenase treatment (Stem Cell Technologies). Generation of iPSC lines were generated with patient consent, and this study was ethically approved (CEN/11/12/066 Central Regional Ethics Committee, New Zealand and 115–16 Ethical Committee LMU Munich, Germany).

Reprogramming of fibroblasts. 2.5 x 10^{-5} NuFF3-RQ human newborn foreskin feeder fibroblasts (GSC-3404, GlobalStem) were seeded per well of a 6-well tissue culture dish with advanced MEM (12491015, Thermo Fisher Scientific) supplemented with 5% HyClone FBS (SV30160.03HI, GE Healthcare), 1% MEM NEAA and GlutaMAX (11140050; 35050061 Thermo Fisher Scientific). Patient fibroblasts were obtained from skin biopsies. Day 1, patients and control fibroblast cultures (CTRL-2522, ATCC) of 70–80% confluency were dissociated using 0.25% trypsin-EDTA (25200056, Life Technologies), counted and seeded on the NuFF3-RQ cells at two different densities: 2 x 10^{-4} cells/well and 4 x 10^{-4} cells/well. Day 2, medium was changed to Pluriton Reprogramming Medium (00–0070, Stemgent) supplemented with 500 ng/ml carrier-free B18R Recombinant Protein (03–0017, Stemgent). Days 3–18, modified mRNA (mmRNA) cocktail was transfected daily combining *OCT4*,*SOX2*, *KLF4*, *LIN28* and *C-MYC* mmRNAs at a 3:1:1:1:1 stoichiometric ratio and Opti-MEM I Reduced Serum Medium (13778150, Thermo Fisher Scientific) in a total volume of 105 µL with a mix of 92 µL Opti-MEM I Reduced Serum Medium and 13 µL Lipofectamine RNAiMAX Transfection Reagent (31985062, Thermo Fisher Scientific) after incubation at room temperature (RT) for 15 min. Cells were transfected for 4 h, then washed, and fresh reprogramming medium supplemented with B18R was added to the cultures. The mmRNA factors were provided by the RNA CORE of the Houston Methodist Hospital and contained the following modifications: 5-methyl CTP, Pseudo-UTP, ARCA cap and a 150-nucleotide poly-A tail. The first morphological changes were noticed as early as day 5 after the first transfection, and the first iPSC colonies appeared by day 12–15. On day 16, medium was changed to STEMPRO hESC SFM (A1000701, Thermo Fisher Scientific) for 5 d. The iPSC colonies were then harvested using 2 mg/ml collagenase, type IV (17104019, Thermo Fisher Scientific) solution in DMEM/F12 (31331093, Thermo Fisher Scientific) by 40- min incubation at 37 °C. The iPSCs were plated on γ-irradiated mouse embryonic fibroblasts (MEFs) and grown in STEMPRO for ten additional passages before adapting the iPSCs to a feeder-free culture system using plates coated with LDEV- Free Geltrex (A1413302, Thermo Fisher Scientific) and mTeSR1 (05850, StemCell Technologies).

CRISPR genome editing for generation of knockout iPSC lines. 409-B2 human iPSCs with a doxycycline-inducible Cas9 nickase (D10A mutation) were incubated with media containing 2 µg/mg doxycycline (Clontech, 631311) 2 d prior to the addition of gRNA by lipofection or electroporation, as described by Riesenberg and Maricic [325]. Genome editing of *DCHS1* and *FAT4* was done by lipofection and electroporation of gRNA (duplex of chemically synthesized crRNA and tracrRNA, alt-CRISPR IDT), respectively. Lipofection by RNAiMAX (Invitrogen, 13778075) was done using a final concentration of 7.5 nM of each gRNA (DCHS1_target1: GTGGACATCAGCATTGTGCC, DCHS1_target2: GGGCACTGGGTT CTGCCTGT). Electroporation was done using the B-16 program of the Nucleofector 2b Device (Lonza) in cuvettes for 100 µL Human Stem Cell nucleofection buffer (Lonza, VVPH-5022) and 0.3 nmol gRNA (FAT4_target1: TTTGATGCTTTCAAAGAAGG, FAT4_target2: GAGATCCTTCTCCGGCAGAG). Edited cells were plated in different wells for analysis and further propagation. 3 d after addition of gRNA, cells for analysis were dissociated using Accutase (SIGMA, A6964), pelleted, resuspended in 15 µL QuickExtract (Epicentre, QE0905T) and incubated at 65 °C for 10 min, 68 °C for 5 min, and finally 98 °C for 5 min. PCR was done in a T100 Thermal Cycler (Bio-Rad) using the KAPA2G Robust PCR Kit (SIGMA, KK5024) with supplied buffer B and 3 µL of cell extract in a total volume of 25 µL. The thermal cycling profile of the PCR was: 95 °C 3 min; 34 x (95 °C 15 s, 65 °C 15 s, 72 °C 15 s); 72 °C 60 s

(DCHS1_forward: GGGTTGTGTGCCTGGACTAT, DCHS1_reverse: TTCCTCTCAGGGCTGTTGAC, FAT4_forward: TAGGGACTGCTGTGCAACTG, FAT4_reverse: AGCTCACAGCCAATCTTCGT). Sample-specific indices on P5 and P7 Illumina adapters were added in a second PCR reaction using Phusion HF MasterMix (Thermo Scientific, F-531L) and 0.3 µL of the first PCR product. The thermal cycling profile of the second PCR was: 98 °C 30 s; 25x (98 °C 10 s, 58 °C 10 s, 72 °C 20 s); 72 °C 5 min. The indexed amplicons were purified using Solid Phase Reversible Immobilization (SPRI) beads [326]. Double-indexed libraries were sequenced on a MiSeq (Illumina), giving paired-end sequences of 2 x 150 bp. After base calling using Bustard (Illumina), adapters were trimmed using leeHom [261], and sequences were analyzed using SAMtools [286]. After confirmed bulk editing success, cells were plated in a single-cell dilution that gave rise to single-cell-derived colonies. DNA isolation and further genotyping of these colonies was done as described above.

DCHS1 genotypes:

This is the amplicon sequence of the *DCHS1* wild type:

wild type:

GGGTTGTGTGCCTGGACTATGCCCACACTGGTGCCTGGTGCCACATC CTCTGGCACAGA AAAAACATACTGTAGTTGCTCAAATATGGGTGGTGT GGGGGTTCCAGGCACAATGCTGATGTCCACTCGGGCACTGGGTTCTGC CTGTAGGCCACCTCCGTCCTCAGCCCCGATCTCCAG CTGCACCACAGAAT TGGCCCGTCTGGCCAAGGGCCAGGCTACTGTCAACAGCCCTGAGAGGAA

3’5’ frame 1

F L S G L L T V A W P L A R R A N S V V Q L E I G A E D G G G L Q A E P S A R V D I S I V P GT P T P P I F E Q L Q Y V F S V P E D V A P GT S V GI V Q A H N

These are the amplicon sequences of the DCHS1 KO:

first allele,

GGGTTGTGTGCCTGGACTATGCCCACACTGGTGCCTGGTGCCACATC CTCTGGCACAGAA AAAACATACTGTAGTTGCTCAAATATGGGTGGTGT GGGGGTTCCAGGCACTGGGTT CTGCCTGTAGGCCACCTCCGTCCTCAG CCCCGATCTCCAGCTGCACCACAGAATTGGCCCGTCTGGCCAAGGGCCA GGCTACTGTCAACAGCCCTGAGAGGAA

3’5’ frame 1

F L S G L L T V A W P L A R R A N S V V Q L E I G A E D G G G L Q A E P S A W N P H T T H I Stop A T T V C F F C A R G C G T R H Q C G H S P G T Q P

second allele,

GGGTTGTGTGCCTGGACTATGCCCACACTGGTGCCTGGTGCCACATC CTCTGGCACAGAAAAAACATACTGTAGTTGCTCAAATATGGGTGGTGT GGGGGTTCCAGGCACAAACTGGG TTCTGCCTGTAGGCCACCTCCGTCC TCAGCCCCGATCTCCAGCTGCACCACAGAATTGGCCCGTCT

GGCCAAGG GCCAGGCTACTGTCAACAGCCCTGAGAGGAA

3’5’ frame 1

F L S G L L T V A W P L A R R A N S V V Q L E I G A E D G G G LQ A E P S L C L E P P H H P Y L S N Y S Met F F L C Q R Met W H Q A P V W A Stop S R H T T

FAT4 genotypes

FAT4 WT (wild type)

4,314 reads

TAGGGACTGCTGTGCAACTGTACAGTGCATATGAAGAGAAC AATAGAACGTTTCTTTTGGCAGCTGTGAAGCGAAATCATAATCA GTATGTGAATCCCAGTGGCGTAGCCACCTTCTTTG AAAGCATCA AAGAGATCCTTCTCCGGCAGAGTGGAGTAAAGGTGGAATCTGTG GATCATGACTCCTGTGTGCATGGCCCATGTCAGAATGGAGGGAG CTGTCTACGAAGATTGGCTGTGAGCT

5'3' frame 3

G T A V Q L Y S A Y E E N N R T F L L A A V K R N H N Q Y V N PS G V A T F F E S I K E I L L R Q S G V K V E S V D H D S C V H G P C Q N G G S C LR R L A VS

FAT4-KO2 (knockout, two different knockout alleles)

6,076 reads

TAGGGACTGCTGTGCAACTGTACAGTGCATATGAAGAGAACAATAGA ACGTTTCTTTTGGCAGCTGTGAAGCGAAATCATAATCAGTATGTGAAT CCCAGTGGCGTAGCCACCTTCTCCGGCAGAGTGGAGTAAAGGTGGAAT CTGTGGATCATGACTCCTGTGTGCATGGCCCATGTCAGAATGGAGGGAG CTGTCTACGAAGATTGGCTGTGAGCT

5'3' frame 3

G T A V Q L Y S A Y E E N N R T F L L A A V K R N H N Q Y V N PS G V A T F S G R V E Stop R W N L W I Met T P V C Met A H V R Met E G A V Y E D W L Stop A

5,345 reads

TAGGGACTGCTGTGCAACTGTACAGTGCATATGAAGAGAACAATAGA ACGTTTCTTTTGGCAGCTGTGAAGCGAAATCATAATCAGTATGTGAAT CCCAGTGGCGTAGCCACCTTCTTTCCGGCAGAGTGGAGTAAAGGTGGA ATCTGTGGATCATGACTCCTGTGTGCATGGCCCATGTCAGAATGGAGGGA GCTGTCTACGAAGATTGGCTGTGAGCT

5'3' frame 3

G T A V Q L Y S A Y E E N N R T F L L A A V K R N H N Q Y VN P S G V A T F F P A E W S K G G I C G S Stop L L C A W P Met S E W R E L S T K I G C E

FAT4-KO3 (knockout, two different knockout alleles)

4,878 reads

TAGGGACTGCTGTGCAACTGTACAGTGCATATGAAGAGAACAATAGA ACGTTTCTTTTGGCAGCTGTGAAGCGAAATCATAATCAGTATGTGAAT CCCAGTGGCGTAGCCACCTTCTCCGGCAGAGTGGAGTAAAGGTGGAAT CTGTGGATCATGACTCCTGTGTGCATGGCCCATGTCAGAATGGAGGGAG CTGTCTACGAAGATTGGCTGTGAGCT

5'3' frame 3

G T A V Q L Y S A Y E E N N R T F L L A A V K R N H N Q Y V N P S G V A T F S G R V E Stop R W N L W I Met T P V C Met A H V R Met E G A V Y E D W L Stop A

4,396 reads

TAGGGACTGCTGTGCAACTGTACAGTGCATATGAAGAGAACAATAGA ACGTTTCTTTTGGCAGCTGTGAAGC
GAAATCATAATCAGTATGTGAAT CCCAGTGGCGTAGCCACCTTCTTTGAAAGCATCAAGGCAGAGTGGAGTA AAGGTGGAATCTGTGGATCATGACTCCTGTGTGCATGGCCCATGTCAGAA TGGAGGGAGCTGTCTACGAAGATTGGCTGTGAGCT

5'3' frame 3

G T A V Q L Y S A Y E E N N R T F L L A A V K R N H N Q Y V N P S G V A T F F E S I K A E W S K G G I C G S Stop L L C A W P Met S E W R E L S T K IG C E

Generation of microRNA against DCHS1 and FAT4. Three miRNA targeting *FAT4* or *DCHS1*, respectively, were cloned into a pENTR entry vector (BlockIT, Invitrogen, Thermo Fisher) and then recombined into a PCAGGS destination vector according to the manufacturer's protocol (BlockIT, Invitrogen, Thermo Fisher). For validation, vectors were electroporated into SH-SY5Y cells and expression of *FAT4* and *DCHS1* quantified using qPCR.

The sequences of the oligos chosen for experiments are:

DCHS1: TGCTGTACACTGTCAGGTTGATCTCCGTTTTGGCCACTGA CTGACGGAGATC ACTGACAG

FAT4: GCTGATCAGTTGCAGTAACAGAGGAGTTTTGGCCACTGACT GACTCCTCTGTCTGCAACT

The qPCR primers sequences used for validation are:

DCHS1-fw: 5' TGCACCTGAAGACACGGTAT 3'
DCHS1-rev: 5' CAGAGGCCTCATAAGCCGTA 3'
FAT4-fw: 5' CTTCCAAATGGACCCTGAGA 3'
FAT4-rev: 5' CGGTGCCCACTTGAGCATTC 3'

Generation of NPCs and neurons. Neural progenitors were generated as described previously [327] with modifications. In short, embryoid bodies were generated from iPSCs by plating colonies in suspension in neural induction medium consisting of DMEM F12 with N2 and B27 supplements (Thermo Fisher). Embryoid bodies were plated on polyornithine and laminin (Sigma Aldrich, St. Louis, MO, USA) coated dishes and cultured for 7 d in neural induction medium. Neural rosettes were manually picked using a stereological microscope (Nikon, Tokyo, Japan) and a P1000 tip, manually dissociated and further cultivated in neural progenitor medium (neural induction medium supplemented with bFGF at 20 ng/ml, Peprotech, Rocky Hill, NJ, USA). For passaging, the cells were dissociated using Accutase (Stem Cell Technologies) and split at a maximum ratio of 1:3.

Generation and analysis of cerebral organoids. Cerebral organoids were generated as previously described [92]. Organoids were kept in 10-cm dishes on an orbital shaker at 37 °C, 5% CO_2 and ambient oxygen level with medium changes every 3 to 4 d. Organoids were analyzed at 20 d, 42 d and 70 d after plating. For immunostaining, 16 µm sections of organoids were prepared using a cryotome. At least five different organoids for each of three different batches were analyzed.

Electroporation of organoids. For electroporation, organoids were kept in neural differentiation medium as described [319] without antibiotics. The organoids were placed in an electroporation chamber (Harvard Apparatus, Holliston, MA, USA), and plasmid DNA was injected at a concentration of 1 µg/µL at several positions. The organoids were then subjected to five pulses at 80 V with a 50-ms duration in an interval of 500 ms using an ECM830 electroporation device (Harvard Apparatus).

3D time-lapse imaging. For 3D time-lapse imaging, slices of electroporated cerebral organoids were prepared and imaged as described previously [328]. Briefly, organoids were sliced at 300 µm thickness on a vibratome (Leica VT1200S) in ice-cold DMEM/F12 (Invitrogen) supplement with sodium bicarbonate, glucose and 10% antibiotics, oxygenated with 100% O_2 for 20 min before cutting. The slices were placed on a cell culture insert (Millicell) and further cultured in normal organoid medium. The slices were kept in an atmosphere with 5% CO_2 at 37 °C. Live imaging was performed for 48 h using Leica TCS SP8 Confocal microscope (Leica, Germany), taking an image every twenty minutes. The cell movement was tracked using ImageJ software and the Manual Tracking Plugin, and the movement parameters calculated and analyzed in GraphPad Prism.

2D time-lapse imaging. For 2D time-lapse imaging, cultures of neural progenitor cells were kept in neural progenitor medium without bFGF for 7 d to allow cells to differentiate. Consequently, live imaging was performed for 3 d using Zeiss Axiovision Observer Fluorescent Microscope (Zeiss, Germany) and Zen Software, taking an image every 5 min. The cell movement was tracked using ImageJ software and the Manual Tracking Plugin, and the movement parameters calculated and analyzed using R. Neuronal identity was confirmed using TUBB3 immunostaining at the end of the imaging experiment.

Dissociation of organoids and neuronal cells in 2D culture for scRNA-seq. Organoids were microdissected to enrich for cortical tissue and washed twice with 2 ml of HBSS ($-Ca^{2+}$/ $-Mg^{2+}$ (w/o), Sigma). The tissue was dissociated using a papain-based neural tissue dissociation kit (Miltenyi) with multiple intermittent triturations using wide-bore pipette tips and p1,000 and p200 pipettes. After dissociation, cells were filtered through a 30 µm strainer and a 20 µm strainer, washed with HBSS ($+Ca^{2+}$/$+Mg^{2+}$ (w), Sigma) and spun down at 300g for

10 min. Cells were then washed again with HBSS (w), spun down for 5 min at 300g and were resuspended in HBSS (w). Counting and viability assessment were performed using Trypan blue staining (Countess automatic cell counter, Invitrogen) for further FACS-based single-cell transcriptomics.

2D differentiated neurons from *DCHS1*, *FAT4* and two controls were detached from wells of a 24-well plate on day 14 after initiation of differentiation into neurons. First, cells were washed with PBS (w/o, Gibco) twice and incubated with 0.5 mM EDTA (Invitrogen) in DPBS (Gibco) + 1% Accutase (v/v, Sigma) for 5 min at 37 °C to detach cells. After incubation, HBSS (w/o, Sigma) was added to the wells, and cells were washed off by rinsing with a 1,000p. Detached cells were spun down 5 min at 300g and washed with HBSS (w/o), spun down again and resuspended in HBSS (w/o) for counting.

RNA-seq experiments. Organoids were microdissected to enrich for cortical regions, and the tissue was dissociated using a papain-based neural tissue dissociation kit (Miltenyi). Neuronal cells in 2D culture were brought into single-cell suspension using 1% (v/v) Accutase (Sigma) in DPBS/EDTA (0.5 mM, Gibco, Invitrogen). Single-cell transcriptomic analysis was performed using either the Smart-seq2 protocol [195] (for all organoid experiments), or the 10X genomics chromium platform (for neuronal cells in 2D differentiation). Libraries were prepared using the Illumina Nextera XT DNA Sample Preparation kit (SS2) or according to the 10X genomics Single Cell 3' v2 protocol. Quantification and quality control of the resulting cDNA was performed using high-throughput capillary gel electrophoresis (Fragment analyzer, Advanced Analytical; for SS2) or the Bioanalyzer High Sensitivity DNA Assay (Agilent; for 10X cDNA, for SS2 and 10X libraries). Single-cell cDNA libraries were sequenced on an Illumina HiSeq2500 platform. Reads were processed and aligned to the genome (hg38 sourced from Ensembl or 10X genomics) using TopHat [329] or using the 10X Genomics Cell Ranger software (v 2.1.0), respectively. Data analysis was performed using custom-written R scripts and using the following packages in R: Seurat (v 2.3.0), Monocle2 (v2.0.0 for SS2 data, v 2.6.3 for 10X data), igraph, ggplot2 and FactoMineR.

For Smart-seq2-based scRNA-seq (organoid scRNA-seq dataset), single cells were isolated by FACS of individual cells into individual wells of 96-well plates. cDNA was prepared from each single cell using the previously described Smart-seq2 protocol [195]. Size distribution and concentration of single-cell cDNA was assessed by high-throughput capillary gel electrophoresis (Fragment analyzer, Advanced Analytical). Sequencing libraries were constructed in 96-well plates using the Illumina Nextera XT DNA Sample Preparation kit as described previously [161]. Libraries were quantified by Agilent Bioanalyzer using the High Sensitivity DNA analysis kit. Up to 192 single-cell libraries were pooled and each cell was sequenced 100-bp paired-end on Illumina HiSeq2500. Base calling, adaptor trimming, and demultiplexing was performed as described in refs. [260, 262]. For 10X Genomics Chromium based scRNA-seq (2D neuronal differentiation dataset), the 10X Chromium Single Cell 3' Kit v2 was used. Control and mutant cells were mixed for multiplexing (*FAT4*/ CTRL and *DCHS1*/ CTRL) and were loaded on one lane of a 10X microfluidic chip. cDNA cleanup, amplification and library preparation were performed by following the Single Cell 3' protocol. Quantification and quality control of the resulting libraries was performed using Bioanalyzer (Agilent) High Sensitivity DNA Assay. The resulting two libraries were pooled at equal ratios and sequenced on two lanes of an Illumina HiSeq2500 platform.

Processing, analysis, and graphic display of Smart-seq2-based scRNA-seq data. Raw reads were processed using a custom script and aligned to a Bowtie2 (ref. [330]) indexed human genome (hg38 sourced from Ensembl) using TopHat [329] with default settings. Transcript levels were quantified as Fragments Per Kilobase of Mapped reads (FPKM) generated by Cufflinks [331] using gencode protein coding genes (hg38 Havana). We excluded cells that had less than 100,000 reads or expressed less than 1,000 genes. Transcript levels were converted to the log-space by taking the $\log_2$ (FPKM). R studio (https://www.rstudio.com/) was used to run custom R scripts to perform principal component analysis (PCA, FactoMineR package), hierarchical clustering (stats package), covariance analysis and to construct heat maps, violin plots, scatter plots and dendrograms. Generally, ggplot2 and gplots packages were used to generate data graphs. The Monocle2 package [323] (https://bioconductor.org/packages/release/bioc/html/monocle.html) was employed to analyze cell lineage relationships using 289 genes (Supplementary Table 3) identified via PCA on all cells and on all neurons. Covariance network analysis and visualizations were done using igraph implemented in R (http://igraph.sf.net). Gene ontology enrichment analyses were performed using DAVID informatics Resources 6.7 (ref. [265]). To assign a cortical zone (VZ, iSVZ, oSVZ, CP) to each individual cell, we calculated for each single cell the Spearman correlation of its transcriptome (all genes) with bulk transcriptome data from each of four microdissected cortical zones (VZ, iSVZ, oSVZ, CP, mean expression value of each gene across 4 replicates)[31] as described previously [114]. Generally, we performed PCA on variable genes (variance > 0.5) expressed (> 1 FPKM) in more than two cells. For Figs. 2c and 4a, we extracted the genes positively and negatively correlating with PC2 and 3, using an absolute PC loading threshold >0.2 with a maximum of 50 genes per PC to avoid individual PCs swamping the analysis, resulting in 200 genes (Supplementary Tables 1 and 3). These genes were plotted in the heat map. The heat map in Extended Data 3o presents the top 50 genes positively and negatively correlating with PC 1. To construct the intercellular correlation networks in Fig. 2d, we computed a pairwise correlation matrix for all cells and using genes discovered in the PCA analysis on cells from the primary human fetal cortex at gestational week 12 and 13 (Supplementary Table 2) as described [114]. We then generated a weighted adjacency network graph using the graph.adjacency() command in igraph and visualized cells as vertices connected to other cells via edges if the Pearson pairwise correlation between two cells was higher than 0.4. The Fruchterman–Reingold layout was used to plot the network graph.

Processing, analysis, and graphic display of 10X genomics–based scRNA-seq data. For the 10X genomics data, base calling, alignment to the human reference GRCh38 (as provided by 10X) and identification of valid cell barcodes was performed using 10X Genomics Cell Ranger software (v 2.1.0). Genotypes for demultiplexing the pooled cell lines in each library were called from merged scRNA-seq organoid datasets of the respective cell lines (*DCHS1*, *FAT4*, CTRL) using bcftools (v 1.4) mpileup and bfctools call. Genotypes for the respective combinations of pooled cell lines were merged using bcftools merge. Demuxlet [282] was used to assign single cells to their respective cell line, taking only valid barcodes into account as defined by Cell Ranger. Cells with insufficient likelihood difference or ALPHA value >0.3 were considered as doublets, and all remaining cells were assigned with their first best cell line assignment as provided by demuxlet. After demultiplexing, most of the cells in the *FAT4*/ CTRL dataset were assigned as CTRL cells, whereas most of the cells in the *DCHS1*/ CTRL dataset were assigned as *DCHS1* cells, which was consistent with the number of cells per line in the cell suspensions loaded on the 10X microfluidic chip device. Thus, we focused the analysis on comparing sufficient numbers of cells of *DCHS1* and CTRL.

Single-cell data were analyzed using the Seurat (v 2.3.0) toolkit for prefiltering and clustering of data. Cells with less than 500 and more than 6,000 genes per cell and more than 10% mitochondrial genes expressed were filtered from the dataset. The normalized data were scaled, and the effect of mitochondrial gene expression and number of UMI's were regressed out. Identification of clusters and *t*-Stochastic Neighbor Embedding (*t*-SNE) clustering was performed using the first ten principal components based on the highly variable genes in the dataset. Mesenchymal-like cells as revealed by clustering and high expression of DCN and COL1A2 and cells with high cycling gene expression (MKI67 and TOP2A $>$ 1.5 UMI after normalizing) were filtered from the data. Cells were ordered along a pseudotemporal trajectory using Monocle (v 2.6.3) by using DDRTree for dimension reduction and using the same ordering genes that were applied to construct differentiation trajectories for organoids.

Immunohistochemistry. Immunostainings were performed as described previously [79]. Nuclei were visualized using 0.1 µg/ml DAPI (Sigma Aldrich). Immunostained sections were analyzed using Olympus or Leica laser-scanning microscopes. F-ACTIN was visualized by incubation with Alexa Fluor 488-conjugated PHALLOIDIN (Thermo Fisher) according to the manufacturer's protocol. Antibodies list is included in the Supplementary Table 5.

FACS analysis. Cerebral organoids (55–60 d in culture) were collected for FACS analysis. Three to six samples were analyzed; every sample contained two individual organoids. Organoids were enzymatically dissociated with Accutase at 37 °C for 30 min. During incubation, every 10 min, the organoids were triturated with a P1000 pipette. After dissociation, samples were washed in PBS via centrifugation at 1,200 r.p.m. for 5 min. The cell suspension was filtered through a 100 µm cell strainer and centrifuged at 1,200 r.p.m. for 5 min, and the cells were fixed in 70% ice-cold ethanol. After 1 h at −20 °C, samples were centrifuged for 30 min at 2,000 r.p.m. and then resuspended in 5 ml staining solution (PBS containing 1% FCS). After further centrifugation for 30 min at 2,000 r.p.m., the cell pellet was resuspended in staining solution containing anti KI67 antibody (1:200, see antibody's table), or anti DCX antibody (1:1,000, see Supplementary Table 5) and incubated for 30 minutes. After washing in staining solution, cells were resuspended in staining solution containing AlexaFluor546 anti–guinea pig, for DCX, or anti-mouse (1:800), for KI67, secondary antibodies and incubated for 30 min. After washing in PBS, cells were resuspended in PBS. FACS analysis was performed on a FACS Aria (BD) in BD FACS Flow TM medium, with a nozzle diameter of 100 µm. Debris and aggregated cells were gated out by forward scatter and sideward scatter; single cells were gated out by FSC-W/FSC-A. Gating for fluorophores was done using samples stained with secondary antibody only. Flow rate was below 500 events/s.

Western blot. Cells were lysed in lysis buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 10% saccharose in H2O) with protease and phosphatase inhibitors (Roche, Basel, Switzerland), and 20 g of protein was separated by SDS-PAGE with a 12% gel. Proteins were transferred to a nitrocellulose membrane (GE Healthcare, Chalfont St Giles, Buckinghamshire, Great Britain). For detection, membranes were incubated with primary antibodies overnight, then with horseradish peroxidase-labeled secondary antibodies at room temperature, and then treated with ECL Western Blotting Detection solution (Millipore, Billerica, MA, USA) to visualize bands. Bands were quantified using ImageJ software.

In situ hybridization. Probes for in situ hybridization were generated as described in ref. 5. Linearized in situ plasmids were in vitro transcribed using DIG NTP labeling mix, RNA polymerase T7 and Sp6 as well as RNase inhibitor (Roche). In situ mRNA transcript detection was performed according to standard procedures.

Cell and tissue quantifications. For quantification of proliferating apical NPCs, all apically located PH3-positive cells were quantified. Cell counts were normalized by the length of the apical surface. MAP2 and ARL13B fluorescence intensity was quantified using ImageJ software. MAP2 signal was quantified only where progenitors are enriched (ventricular zone). ARL13B signal was quantified at the apical surface and at basal positions. In particular, the ARL13B+ fluorescent signal intensity from all the cilia facing the ventricular lumen versus the rest of the cilia found in the germinal and cortical zones was measured using ImageJ software and normalized per area. The analysis of disorganization of the ventricles in mutant and KO lines was done based on the MAP2 staining. More precisely as heterotopic were called the ventricles, which have clusters of MAP2-positive cells in the germinal zones, and as disorganized were defined the ventricles in which the boundaries between the cortical and germinal zone was not clearly distinguished. The percentage of ventricles with NEUN+ cells only in the cortical plate or having more than five NEUN+ cells the germinal zone in addition to the normal NEUN layer in the cortical plate was quantified. The apical surface was identified using β-CATENIN and PALS1 immunostaining, and the apical distance was measured from the apical surface, defined by the immunostainings, to the cell nuclei, defined by the DAPI, in at least five positions per germinal zone using ImageJ software. The organoid size was measured based on the surface of the organoids using ImageJ software. The thickness of the ventricular zone and cortical plate were manually outlined and measured using ImageJ software. NPCs radial morphology and tortuosity were assessed via manual outlining of the NESTIN+ processes or the GFP+ processes after electroporation using ImageJ software. The tortuosity index was calculated by dividing the total (segmented) length of the NESTIN+ of GFP+ processes divided by the straight line connecting the beginning and end of the segmented line. The identity of cells tracked through time lapse in 2D culture was assessed by the expression of MAP2 and DCX.

More details regarding statistical tests applied and individual experiments replicates are found in the figure legends, where b refers to batches, o refers to organoids, v refers to ventricles and c refers to cells.

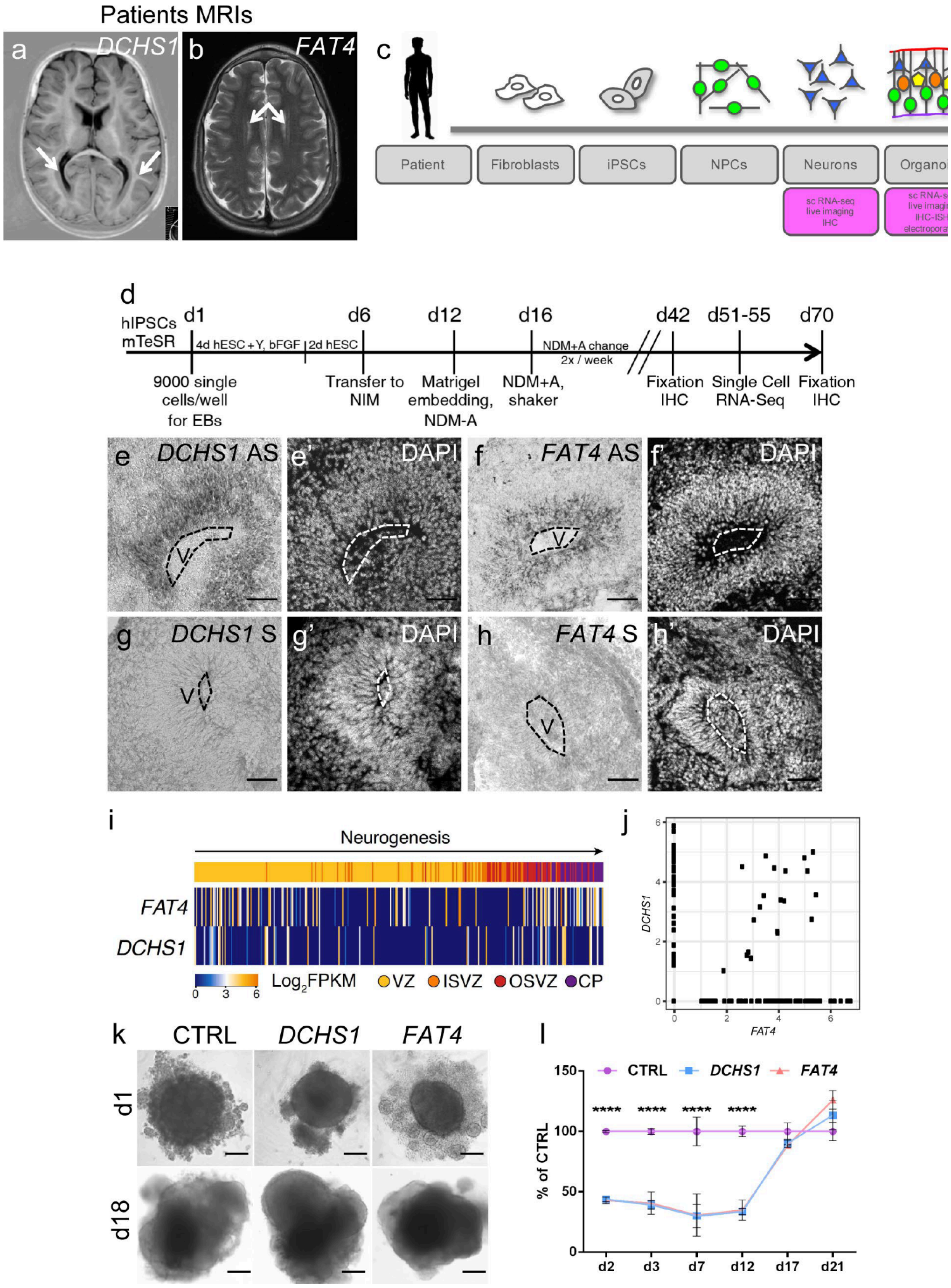
All analysis was done using the R statistical software package, the GraphPad software and SPSS 20.0.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

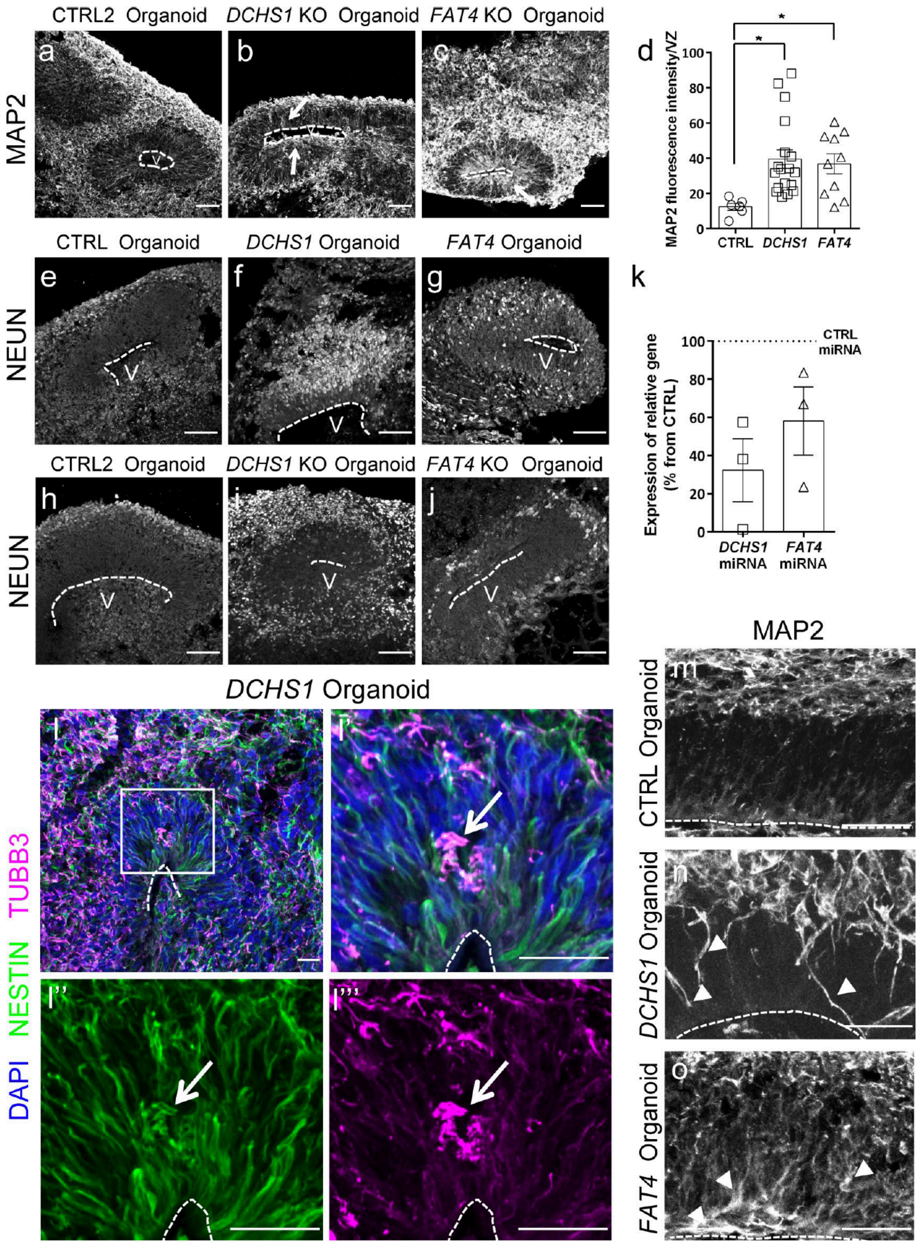
The scRNA-seq data used in this study have been deposited in the Gene Expression Omnibus under accession number GSE124031. All relevant accession codes are provided. Further details on the methods can be found in the Life Sciences Reporting Summary. Additional data that support the findings of this study are available from the corresponding author upon reasonable request.

Extended data

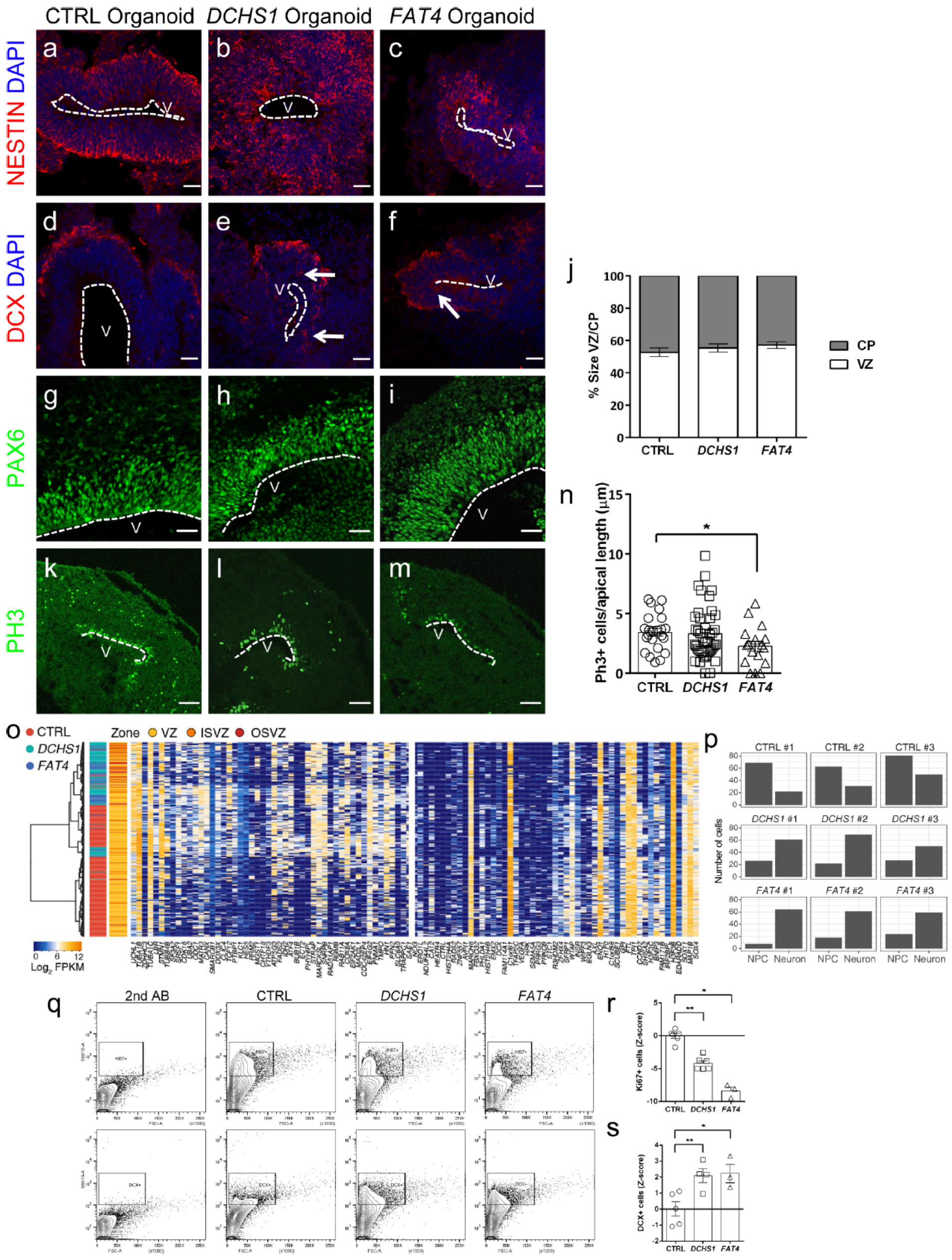


Extended Data Fig. 1: Expression of DCHS1 and FAT4 in cerebral organoids and temporal development of DCHS1- and FAT4-mutant organoids. **a**, Axial T1 image demonstrating laminar PH lining the occipital horns and peritrigonal region of the ventricles (arrows). **b**, Axial T2 demonstrating linear lesions iso-intense with cortical gray matter adjacent to the lateral walls of the bodies of the lateral ventricles [313]. **c**, Schematic representation of the main experimental approach used. **d**, Timeline of the organoid generation. EBs, embryoid bodies; hESC, low-bFGF human embryonic stem cell medium; NIM, neural induction medium; Y, Rock inhibitor; NDM, neural differentiation medium; +/− A: B27 supplement with or without vitamin; IHC, immunohistochemistry; RNA-seq, RNA sequencing.

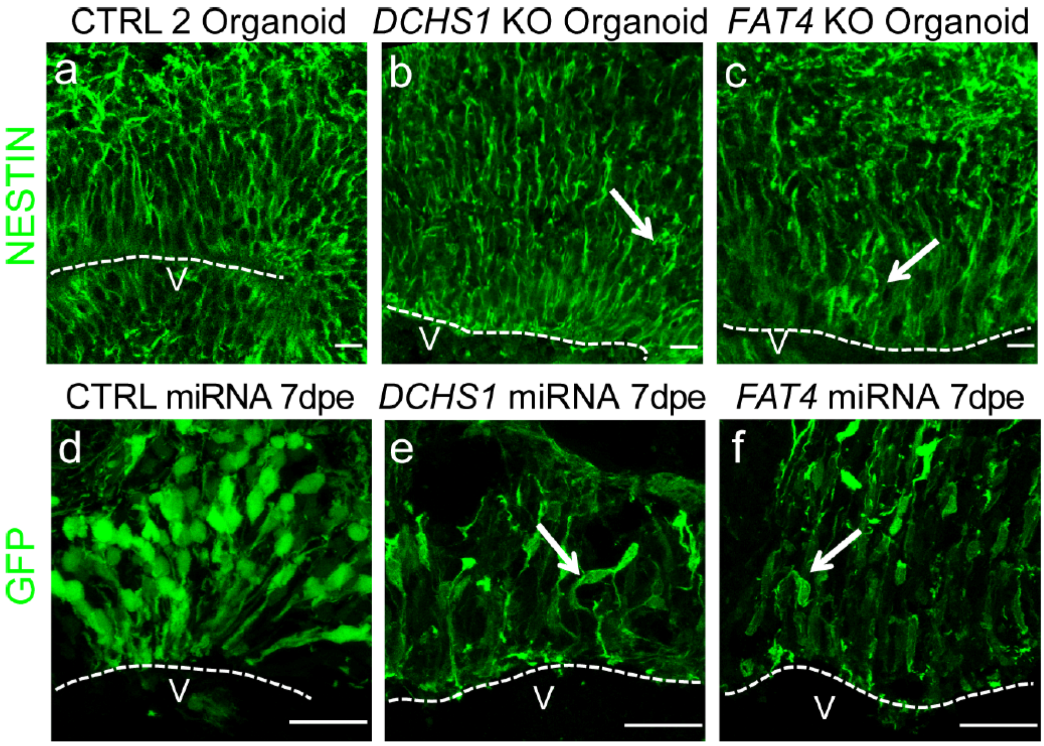
e–h', Detection of DCHS1 and FAT4 by in situ hybridization, $b = 2$, $o = 5$ per condition. **i**,**j**, mRNA expression of *DCHS1* and *FAT4* in single cells ($c = 316$) derived from control organoids, showing similar expression patterns between *DCHS1* and *FAT4* but often in different cells. **i**, Cells (columns) are ordered based on their PC 2 loading, corresponding to the trajectory from NPCs to neurons. Side bar shows maximal zone correlation for each single cell (ventricular zone, VZ, yellow; inner subventricular zone, iSVZ, orange; outer SVZ, oSVZ, red; cortical plate, CP, purple). **j**, Biplot showing transcript levels (in $\log_2$(FPKM)) of *FAT4* (x axis) and *DCHS1* (y axis) in 316 single cells of control organoids. **k**,**l**, Temporal development of patient-derived cerebral organoids compared with control organoids ($o = 14$ CTRL, 17 *DCHS1*, 17 *FAT4*); the diameters of organoids derived from *DCHS1*- and *FAT4*-mutant cells are slightly smaller compared to those of control organoids until day 12 (d12), shown in **k**. Significance based on twoway ANOVA, $P = 0.000$, Tukey HSD post hoc for multiple comparisons was performed for defining statistical differences between the three genotypes. Dotted lines highlight ventricles (V). Data in graphs are represented as mean ± s.e.m. Scale bars, 100 µm in **e–h'**,**k**



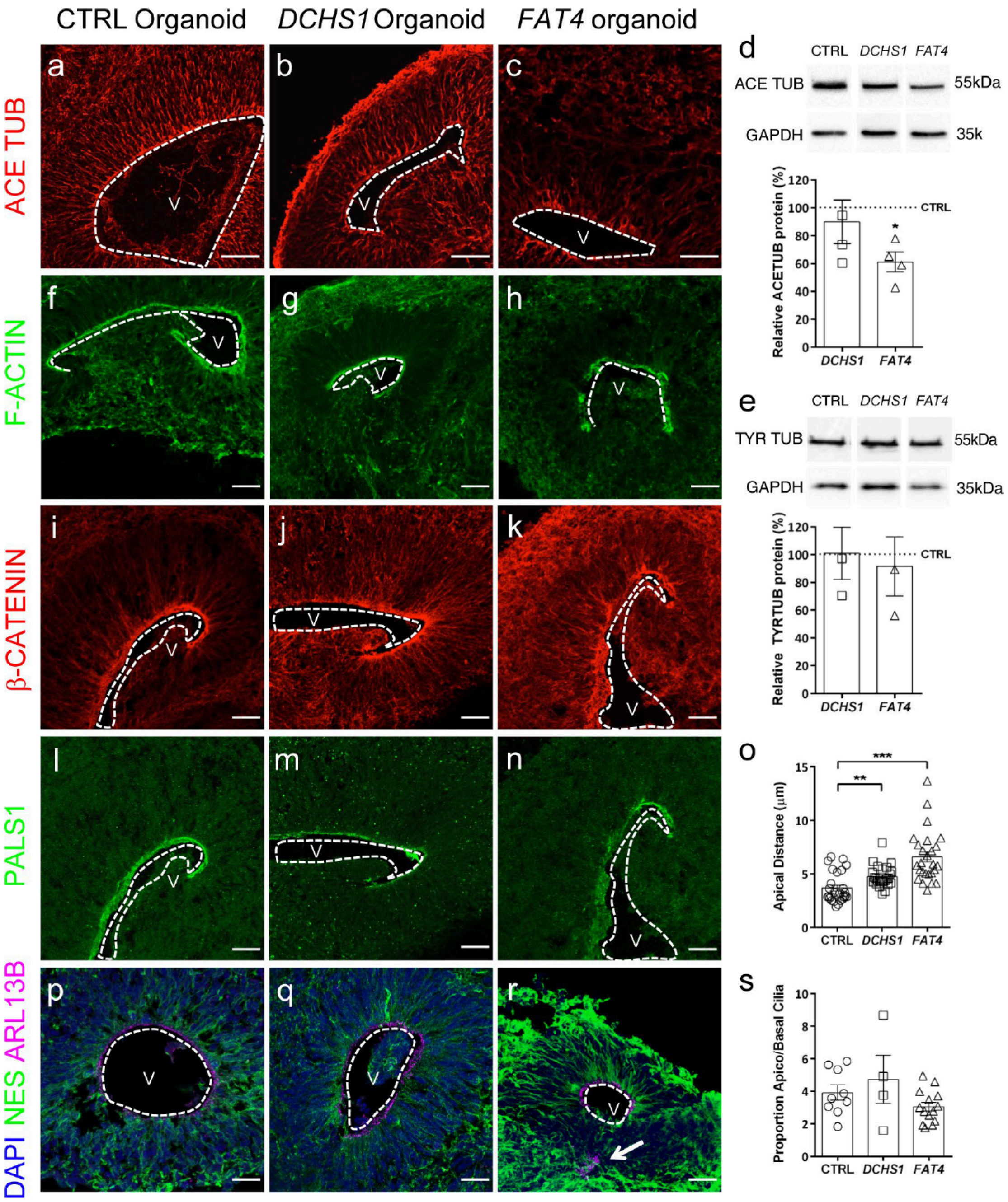
Extended Data Fig. 2: Heterotopically located neurons in DCHS1 and FAT4-mutant and KO organoids. **a–c**,**e–j**,**l–o**, Micrographs of sections of mutant or KO organoids immunostained as indicated in the panels. Note the mispositioning of neurons marked by arrows in mutant or in electroporated cerebral organoids. **a–c**, $b = 2$, $o = 6$ per condition; **e–j**, $b = 2$, $o = 6$ per condition; **l–l''**', $b = 5$, $o = 15$ per condition; **m–o**, $b = 5$, $o = 15$ per condition. **d**, MAP2 fluorescence intensity measured only in the ventricular zones (VZ) (v = 6 CTRL, 18 *DCHS1*, 11 *FAT4*; significance based on one-way ANOVA, $P = 0.0166$, Tukey HSD post hoc for multiple comparisons for defining statistical differences between the three genotypes). **k**, Quantification measured by qPCR of the knockdown of *DCHS1* and *FAT4* by microRNAs (miRNA) against *DCHS1* or *FAT4*, respectively ($b = 2$, independent cultures per time = 3 CTRL, 3 miRNA *DCHS1*, 3 miRNA *FAT4*, significance based on one-way ANOVA, $P = 0.0377$, Tukey HSD post hoc for multiple comparisons for defining statistical differences between the three genotypes) in SH-SY5Y cells 48 h after nucleofection. **l–l''**', Nodule of TUBB3+ neurons intermingling with NESTIN+ processes of NPCs in the germinal zone of *DCHS1*-mutant organoids. **m–o** *DCHS1*- and *FAT4*-mutant organoids show changes in the morphology and thickness of their neuritis, as depicted by arrowheads. Dotted lines highlight ventricles (V). Data in graphs are represented as mean ± s.e.m. Scale bars, 100 μm in **a–c** and **e–j** and 30 μm in **l–o**. Source data.



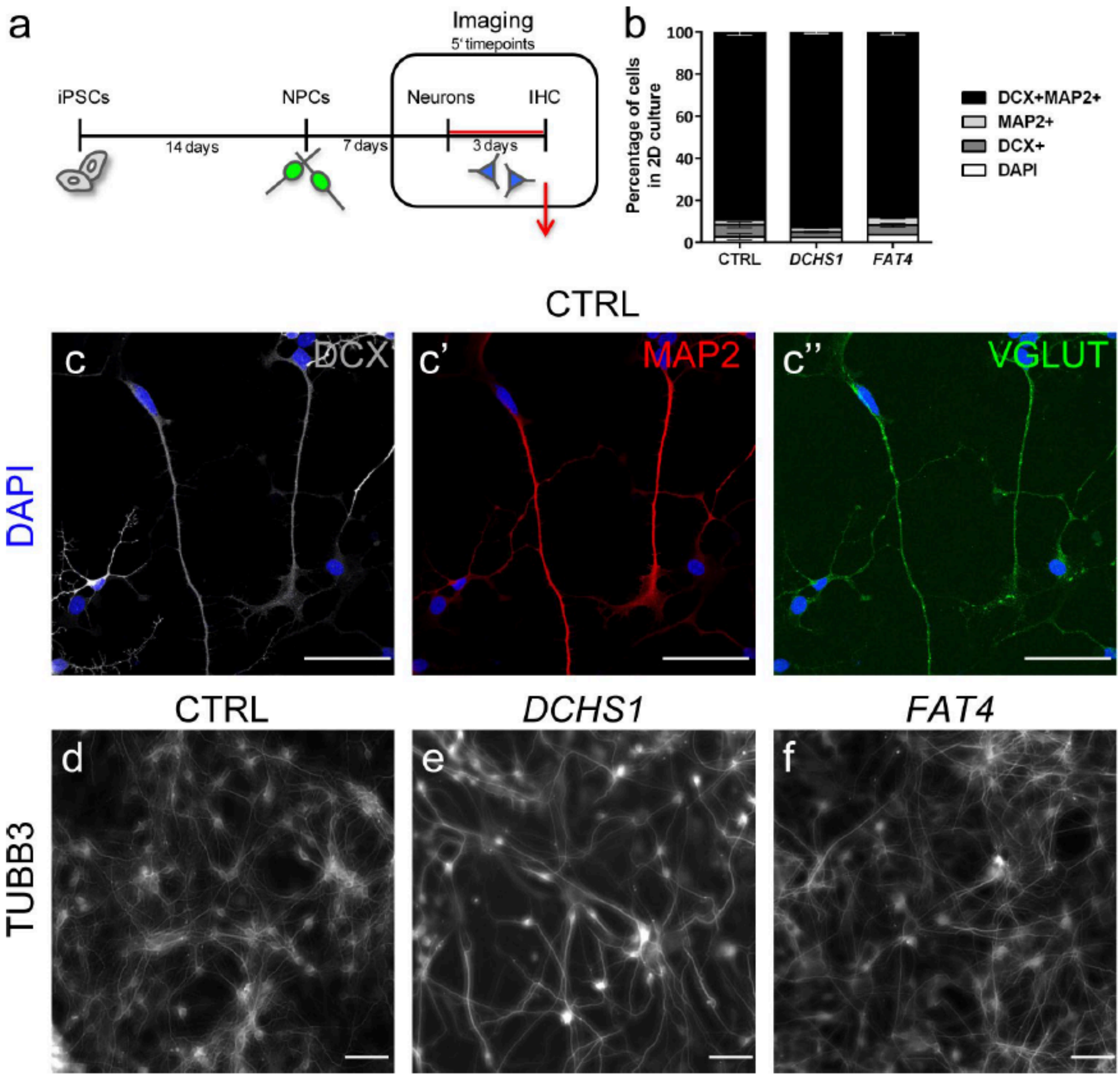
Extended Data Fig. 3: Neural progenitor proliferation and signatures in mutant organoids. **a–i**,**k–m**, Micrographs of sections of mutant cerebral organoids from day 20 immunostained for NESTIN (**a–c**) and DCX (**d–f**) and from day 42 immunostained for PAX6 (**g–i**) and PH3 (**k–m**). **a–f**, $b = 1$, $o = 3$ per condition; **g–i**,**k–m**, $b = 3$, $o = 9$ per condition. **j**, Quantification of thickness of ventricular zone (VZ) and cortical plate (CP) structures in cerebral organoids ($v = 6$ CTRL, 11 *DCHS1*, 12 *FAT4*). n, Quantification of PH3+ cells per length of apical surface ($o = 3$, $v = 23$ CTRL, 41 *DCHS1*, 17 *FAT4*, $F(2,80) = 2.41$, $P = 0.097$, comparison between *CTRL* and *FAT4* $F(1,39) = 5.18$, $P = 0.029$). **o**, Hierarchical clustering visualizing for all NPCs (338 single cells), expression of genes identified by PCA (top 50 positively and negatively correlating with PC 1) on all NPCs. **p**, Number of NPCs and neurons for each experiment shown in **o**. **q**, FACS plots depicting the definition of the sorting gates (secondary antibodies control) and sorting of KI67+ or DCX+ cells in control, *DCHS1*- and *FAT4*-mutant organoids ($b = 2$, $o = 6$ CTRL, 6 *DCHS1*, 3 *FAT4*). **r**,**s**, Z scores of the quantification of KI67+ cells (**r**) and DCX+ cells (**s**) from FACs analysis shown in **q**. Statistical analysis was performed using two-tailed Mann Whitney test. **r**, CTRL to *DCHS1* $P = 0.0022$, CTRL to *FAT4* $P = 0.0238$. **s**, CTRL to *DCHS1* $P = 0.0317$, CTRL to *FAT4* $P = 0.0357$. Results are mean ± s.e.m. (j,n) or z scores as mean ± s.e.m. (r,s). Dotted lines highlight ventricles (V). Scale bars, 30 µm for **a-f** and 50 µm for **g-m**.



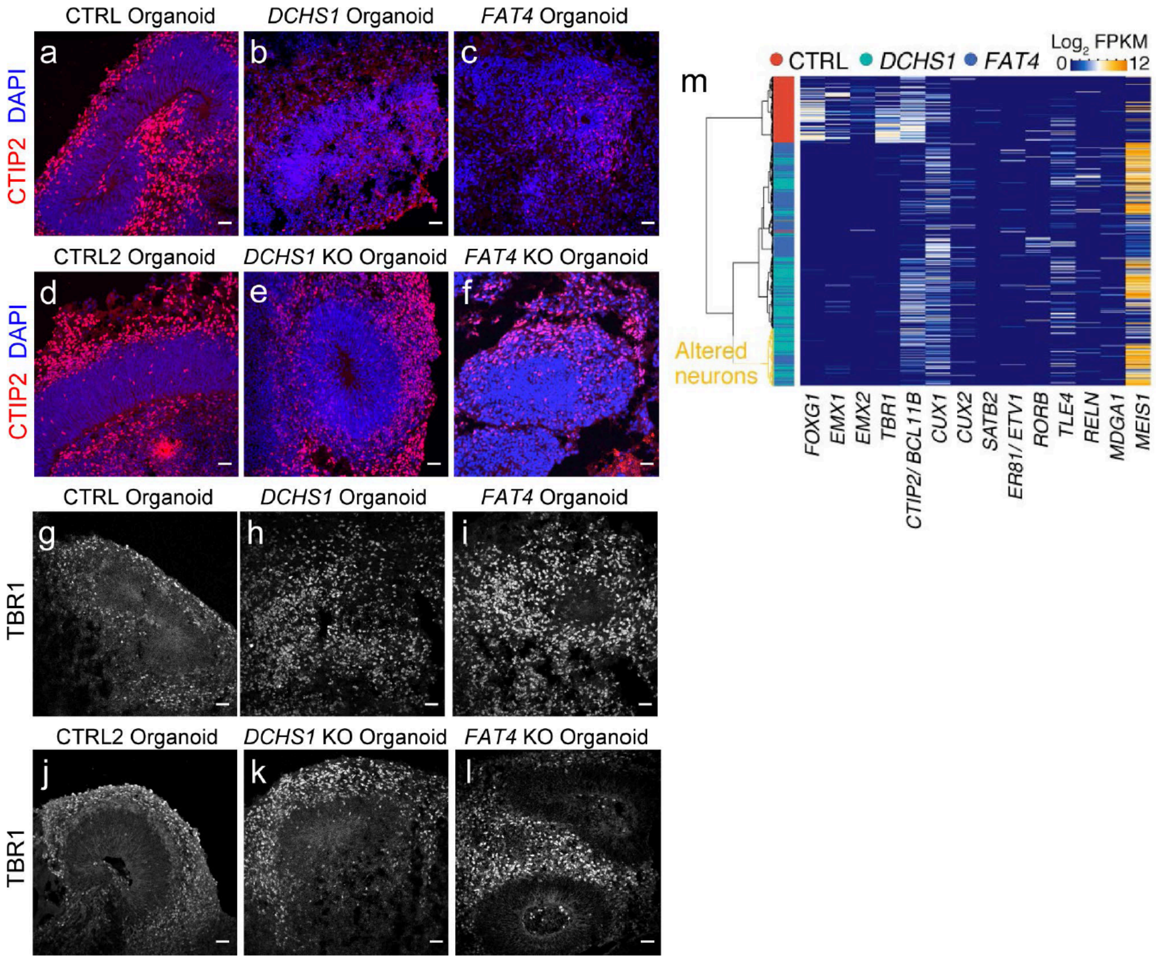
Extended Data Fig. 4: Morphological changes in the NPCs upon DCHS1 and FAT4 deletion. **a**-**c**, Micrographs of sections of KO cerebral organoids day 40 immunostained for NESTIN. Arrows indicate the disrupted morphology of NPCs (**b**,**c**). **d**-**f**, Micrographs of sections of control organoids electroporated with miRNA against *DCHS1* or *FAT4* at day 42 and analyzed at day 49. Arrows indicate the disrupted morphology upon downregulation of *DCHS1* and *FAT4* (**e**,**f**). **a**-**c**, $b = 2$, $o = 6$ per condition; **d**-**f**, $b = 2$, $o = 6$ per condition. Dotted lines highlight ventricles (V). Scale bars, 30 µm.



Extended Data Fig. 5: Apicobasal polarity in cerebral organoids. **a**-**c**, Micrographs of sections of organoids (day 42) immunostained for ACETYLATED TUBULIN, $b = 2$, $o = 6$ per condition. d,e, Western blot and quantification of ACETYLATED TUBULIN (d) (independent cultures = 5 CTRL, 4 *DCHS1*, 4 *FAT4*, significance based on one sample two-tailed t test, $P = 0.562$ CTRL vs *DCHS1*, $P = 0.013$ CTRL vs *FAT4*) and TYROSINATED TUBULIN (e) (independent cultures = 3 CTRL, 3 *DCHS1*, 3 *FAT4*) levels in NPCs, significance based on one sample two-tailed t test, $P = 0.967$ CTRL vs *DCHS1*, $P = 0.728$ CTRL vs *FAT4*. **f**-**n**,**p**-**r**, Micrographs of sections of cerebral organoids (day 42) immunostained as indicated in the panels. **f**-**n**,**p**-**r**, $b = 3$, $o = 9$ per condition. **o**, Quantification of the distance from the apical surface (positive for β-CATENIN and PALS1) and DAPI+ nuclei of NPCs (v = 27 CTRL, 26 *DCHS1*, 26 *FAT4*; 5 different positions were measured and averaged for each ventricle; significance based on one-sample two-tailed t test, $P = 0.0025$ CTRL vs *DCHS1*, $P = 0.001$ CTRL vs *FAT4*). **s**, Ratio of ARL13B fluorescence intensity measured at the apical surface (cilia facing the ventricular lumen) and at basal position (all the rest of the cilia in the germinal and cortical zones) ($b = 5$, $o = 15$, $v = 9$ CTRL, 4 *DCHS1*, 14 *FAT4*, significance based on one-sample t test). Results are mean ± s.e.m. Dotted lines highlight ventricles (V). Scale bar, 50 µm (**a**-**c**) and 20 µm (**f**-**r**).



Extended Data Fig. 6: 2D time-lapse imaging experimental design and morphological changes of mutant neurons. **a**, Experimental design for 2D live imaging of migrating neurons. Neural progenitors were differentiated for 7 d and imaged for 3 d every 5 min. **b**, Quantification of the percentage of cells expressing DCX, MAP2 or both markers in neuronal cultures after 10 d in culture. **c–c''**, Immunostaining for DoubleCortin (DCX), MAP2 (mature neurons) and VGLUT (glutamatergic neurons) in 2D neurons derived from control cells in monolayer culture, b = 3, independent cultures = 9 per condition. **d–f**, Immunostaining for TUBB3 in 2D neurons derived from control and mutant cells in monolayer culture, b = 3, independent cultures = 9 per condition. Results are mean ± s.e.m. Scale bar, 20 µm. Nat



Extended Data Fig. 7: Characterization of mutant and KO organoid regions. a-**l**, Micrographs of sections of mutant and KO organoids immunostained as depicted in the panels. **a**-**c**,**g**-**i**, $b = 3$, $o = 9$ per condition; **d**-**f**,**j**-**l**, $b = 2$, $o = 6$ per condition. **m**, Heat map showing expression of genes marking neurons in the cortex/forebrain (columns) for all neuronal cells from control and mutant organoids (rows, 467 single cells). Scale bar, 30 µm.

Supplementary Information

Supplementary information can be found in the online version of the article (doi: 10.1038/s41591-019-0371-0).

Reporting summary

Supplementary tables 1-5

Sourca Data Fig. 1 Statistical Source Data

Sourca Data Fig. 3 Statistical Source Data

Sourca Data Extended Data Fig. 1 Statistical Source Data

Sourca Data Extended Data Fig. 2 Statistical Source Data

Sourca Data Extended Data Fig. 3 Statistical Source Data

Sourca Data Extended Data Fig. 5 Unprocessed Western Blots

Sourca Data Extended Data Fig. 5 Statistical Source Data

Sourca Data Extended Data Fig. 6 Statistical Source Data

References

1. Sherwood, C. C., Subiaul, F. & Zawidzki, T. W. A natural history of the human mind: tracing evolutionary changes in brain and cognition. *J Anat* **212,** 426–54 (2008).
2. Boesch, C. & Boesch, H. Tool use and tool making in wild chimpanzees. *Folia Primatol (Basel)* **54,** 86–99 (1990).
3. Whiten, A. *et al.* Cultures in chimpanzees. *Nature* **399,** 682–5 (1999).
4. Silbereis, J. C., Pochareddy, S., Zhu, Y., Li, M. & Sestan, N. The cellular and molecular landscapes of the developing human central nervous system. *Neuron* **89,** 248–68 (2016).
5. Taverna, E., Götz, M. & Huttner, W. B. The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex. *Annu Rev Cell Dev Biol* **30,** 465–502 (2014).
6. Noctor, S. C. *et al.* Dividing precursor cells of the embryonic cortical ventricular zone have morphological and molecular characteristics of radial glia. *J Neurosci* **22,** 3161–73 (2002).
7. Noctor, S. C., Martinez-Cerdeno, V., Ivic, L. & Kriegstein, A. R. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat Neurosci* **7,** 136–44 (2004).
8. Haubensak, W., Attardo, A., Denk, W. & Huttner, W. B. Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: a major site of neurogenesis. *Proc Natl Acad Sci U S A* **101,** 3196–201 (2004).
9. Stancik, E. K., Navarro-Quiroga, I., Sellke, R. & Haydar, T. F. Heterogeneity in ventricular zone neural precursors contributes to neuronal fate diversity in the postnatal neocortex. *J Neurosci* **30,** 7028–36 (2010).
10. Brown, K. N. *et al.* Clonal production and organization of inhibitory interneurons in the neocortex. *Science* **334,** 480–6 (2011).
11. Ma, T. *et al.* Subcortical origins of human and monkey neocortical interneurons. *Nat Neurosci* **16,** 1588–97 (2013).
12. Hansen, D. V. *et al.* Non-epithelial stem cells and cortical interneuron production in the human ganglionic eminences. *Nat Neurosci* **16,** 1576–87 (2013).

13. Miller, D. J. *et al.* Prolonged myelination in human neocortical evolution. *Proc Natl Acad Sci U S A* **109,** 16480–5 (2012).
14. Herculano-Houzel, S. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proc Natl Acad Sci U S A* **109 Suppl 1,** 10661–8 (2012).
15. Von Bartheld, C. S., Bahney, J. & Herculano-Houzel, S. The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting. *J Comp Neurol* **524,** 3865–3895 (2016).
16. Azevedo, F. A. *et al.* Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J Comp Neurol* **513,** 532–41 (2009).
17. Buchsbaum, I. Y. & Cappello, S. Neuronal migration in the CNS during development and disease: insights from in vivo and in vitro models. *Development* **146** (2019).
18. Willsey, A. J. *et al.* Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155,** 997–1007 (2013).
19. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362** (2018).
20. Pollen, A. A. *et al.* Establishing cerebral organoids as models of human-specific brain evolution. *Cell* **176,** 743–756 e17 (2019).
21. Won, H., Huang, J., Opland, C. K., Hartl, C. L. & Geschwind, D. H. Human evolved regulatory elements modulate genes involved in cortical expansion and neurodevelopmental disease susceptibility. *Nat Commun* **10,** 2396 (2019).
22. Robson, S. L. & Wood, B. Hominin life history: reconstruction and evolution. *J Anat* **212,** 394–425 (2008).
23. Geschwind, D. H. & Rakic, P. Cortical evolution: judge the brain by its cover. *Neuron* **80,** 633–47 (2013).
24. Smart, I. H., Dehay, C., Giroud, P., Berland, M. & Kennedy, H. Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cereb Cortex* **12,** 37–53 (2002).
25. Hansen, D. V., Lui, J. H., Parker, P. R. & Kriegstein, A. R. Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* **464,** 554–561 (2010).
26. Fietz, S. A. *et al.* OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. *Nat Neurosci* **13,** 690–9 (2010).

27. Shitamukai, A., Konno, D. & Matsuzaki, F. Oblique radial glial divisions in the developing mouse neocortex induce self-renewing progenitors outside the germinal zone that resemble primate outer subventricular zone progenitors. *J Neurosci* **31,** 3683–95 (2011).
28. Betizeau, M. *et al.* Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate. *Neuron* **80,** 442–57 (2013).
29. Pollen, A. A. *et al.* Molecular identity of human outer radial glia during cortical development. *Cell* **163,** 55–67 (2015).
30. Kalebic, N. *et al.* Neocortical expansion due to increased proliferation of basal progenitors Is linked to changes in their morphology. *Cell Stem Cell* **24,** 535–550 e9 (2019).
31. Fietz, S. A. *et al.* Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal. *Proc Natl Acad Sci U S A* **109,** 11836–41 (2012).
32. Yang, Y. & Raine, A. Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: a meta-analysis. *Psychiatry Res* **174,** 81–8 (2009).
33. Smaers, J. B., Gomez-Robles, A., Parks, A. N. & Sherwood, C. C. Exceptional evolutionary expansion of prefrontal cortex in great apes and humans. *Curr Biol* **27,** 714–720 (2017).
34. Gabi, M. *et al.* No relative expansion of the number of prefrontal neurons in primate and human evolution. *Proc Natl Acad Sci U S A* **113,** 9617–22 (2016).
35. Semendeferi, K., Lu, A., Schenker, N. & Damasio, H. Humans and great apes share a large frontal cortex. *Nat Neurosci* **5,** 272–6 (2002).
36. Sherwood, C. C. & Smaers, J. B. What's the fuss over human frontal lobe evolution? *Trends Cogn Sci* **17,** 432–3 (2013).
37. Schoenemann, P. T., Sheehan, M. J. & Glotzer, L. D. Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nat Neurosci* **8,** 242–52 (2005).
38. Donahue, C. J., Glasser, M. F., Preuss, T. M., Rilling, J. K. & Van Essen, D. C. Quantitative assessment of prefrontal cortex in humans relative to nonhuman primates. *Proceedings of the National Academy of Sciences of the United States of America* **115,** E5183–E5192 (2018).
39. Gomez-Robles, A., Hopkins, W. D., Schapiro, S. J. & Sherwood, C. C. Relaxed genetic control of cortical organization in human brains compared with chimpanzees. *Proc Natl Acad Sci U S A* **112,** 14799–804 (2015).

40. Semendeferi, K. *et al.* Spatial organization of neurons in the frontal pole sets humans apart from great apes. *Cereb Cortex* **21,** 1485–97 (2011).

41. Bianchi, S. *et al.* Dendritic morphology of pyramidal neurons in the chimpanzee neocortex: regional specializations and comparison to humans. *Cereb Cortex* **23,** 2429–36 (2013).

42. Hutsler, J. J., Lee, D. G. & Porter, K. K. Comparative analysis of cortical layering and supragranular layer enlargement in rodent carnivore and primate species. *Brain Res* **1052,** 71–81 (2005).

43. Nimchinsky, E. A. *et al.* A neuronal morphologic type unique to humans and great apes. *Proc Natl Acad Sci U S A* **96,** 5268–73 (1999).

44. Allman, J. M. *et al.* The von Economo neurons in frontoinsular and anterior cingulate cortex in great apes and humans. *Brain Struct Funct* **214,** 495–517 (2010).

45. Butti, C., Santos, M., Uppal, N. & Hof, P. R. Von Economo neurons: clinical and evolutionary perspectives. *Cortex* **49,** 312–26 (2013).

46. Sousa, A. M. M. *et al.* Molecular and cellular reorganization of neural circuits in the human lineage. *Science* **358,** 1027–1032 (2017).

47. Boldog, E. *et al.* Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type. *Nat Neurosci* **21,** 1185–1195 (2018).

48. Oberheim, N. A. *et al.* Uniquely hominid features of adult human astrocytes. *J Neurosci* **29,** 3276–87 (2009).

49. Sherwood, C. C. *et al.* Evolution of increased glia-neuron ratios in the human frontal cortex. *Proc Natl Acad Sci U S A* **103,** 13606–11 (2006).

50. Fu, X. *et al.* Rapid metabolic evolution in human prefrontal cortex. *Proc Natl Acad Sci U S A* **108,** 6181–6 (2011).

51. Bozek, K. *et al.* Exceptional evolutionary divergence of human muscle and brain metabolomes parallels human cognitive and physical uniqueness. *PLoS Biol* **12,** e1001871 (2014).

52. Bozek, K. *et al.* Organization and evolution of brain lipidome revealed by large-scale analysis of human, chimpanzee, macaque, and mouse tissues. *Neuron* **85,** 695–702 (2015).

53. Li, Q. *et al.* Changes in lipidome composition during brain development in humans, chimpanzees, and macaque monkeys. *Mol Biol Evol* **34,** 1155–1166 (2017).

54. Walker, R., Hill, K., Burger, O. & Hurtado, A. M. Life in the slow lane revisited: ontogenetic separation between chimpanzees and humans. *Am J Phys Anthropol* **129,** 577–83 (2006).

55. Somel, M. *et al.* Transcriptional neoteny in the human brain. *Proc Natl Acad Sci U S A* **106,** 5743–8 (2009).

56. Liu, X. *et al.* Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques. *Genome Res* **22,** 611–22 (2012).

57. Bakken, T. E. *et al.* A comprehensive transcriptional map of primate brain development. *Nature* **535,** 367–75 (2016).

58. Petanjek, Z. *et al.* Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proc Natl Acad Sci U S A* **108,** 13281–6 (2011).

59. Otani, T., Marchetto, M. C., Gage, F. H., Simons, B. D. & Livesey, F. J. 2D and 3D stem cell models of primate cortical development identify species-specific differences in progenitor behavior contributing to brain size. *Cell Stem Cell* **18,** 467–80 (2016).

60. McLean, C. Y. *et al.* Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature* **471,** 216–9 (2011).

61. Florio, M. *et al.* Evolution and cell-type specificity of human-specific genes preferentially expressed in progenitors of fetal neocortex. *Elife* **7,** 1–37 (2018).

62. Florio, M. *et al.* Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science* **347,** 1465–70 (2015).

63. Fiddes, I. T. *et al.* Human-specific NOTCH2NL genes affect Notch signaling and cortical neurogenesis. *Cell* **173,** 1356–1369 (2018).

64. Suzuki, I. K. *et al.* Human-specific NOTCH2NL genes expand cortical neurogenesis through Delta/Notch regulation. *Cell* **173,** 1370–1384 (2018).

65. Dennis, M. Y. *et al.* Evolution of human-specific neural SRGAP2 genes by incomplete segmental duplication. *Cell* **149,** 912–22 (2012).

66. Charrier, C. *et al.* Inhibition of SRGAP2 function by its human-specific paralogs induces neoteny during spine maturation. *Cell* **149,** 923–35 (2012).

67. Li, Q. *et al.* Mice carrying a human GLUD2 gene recapitulate aspects of human transcriptome and metabolome development. *Proc Natl Acad Sci U S A* **113,** 5358–63 (2016).

68. Enard, W. *et al.* A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell* **137,** 961–71 (2009).

69. Schreiweis, C. *et al.* Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proc Natl Acad Sci U S A* **111,** 14253–8 (2014).

70. Tsui, D., Vessey, J. P., Tomita, H., Kaplan, D. R. & Miller, F. D. FoxP2 regulates neurogenesis during embryonic cortical development. *J Neurosci* **33,** 244–58 (2013).

71. Bae, B. I. *et al.* Evolutionarily dynamic alternative splicing of GPR56 regulates regional cerebral cortical patterning. *Science* **343,** 764–8 (2014).

72. Pollard, K. S. *et al.* An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* **443,** 167–72 (2006).

73. Capra, J. A., Erwin, G. D., McKinsey, G., Rubenstein, J. L. & Pollard, K. S. Many human accelerated regions are developmental enhancers. *Philos Trans R Soc Lond B Biol Sci* **368,** 20130025–1–20130025–13 (2013).

74. Kamm, G. B., Lopez-Leal, R., Lorenzo, J. R. & Franchini, L. F. A fast-evolving human NPAS3 enhancer gained reporter expression in the developing forebrain of transgenic mice. *Philos Trans R Soc Lond B Biol Sci* **368,** 20130019–1–20130019–11 (2013).

75. Boyd, J. L. *et al.* Human-chimpanzee differences in a FZD8 enhancer alter cell-cycle dynamics in the developing neocortex. *Curr Biol* **25,** 772–779 (2015).

76. Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508,** 199–206 (2014).

77. Reillo, I. & Borrell, V. Germinal zones in the developing cerebral cortex of ferret: ontogeny, cell cycle kinetics, and diversity of progenitors. *Cereb Cortex* **22,** 2039–54 (2012).

78. Liao, B. Y. & Zhang, J. Null mutations in human and mouse orthologs frequently result in different phenotypes. *Proc Natl Acad Sci U S A* **105,** 6987–92 (2008).

79. Cappello, S. *et al.* Mutations in genes encoding the cadherin receptor-ligand pair DCHS1 and FAT4 disrupt cerebral cortical development. *Nat Genet* **45,** 1300–8 (2013).

80. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–76 (2006).

81. Okita, K. *et al.* An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. *Stem Cells* **31,** 458–66 (2013).

82. Marchetto, M. C. N. *et al.* Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature* **503,** 525–529 (2013).

83. Gallego Romero, I. *et al.* A panel of induced pluripotent stem cells from chimpanzees: a resource for comparative functional genomics. *Elife* **4,** e07103 (2015).

84. Wunderlich, S. *et al.* Primate iPS cells as tools for evolutionary analyses. *Stem Cell Res* **12,** 622–9 (2014).

85. Marchetto, M. C. *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* **143,** 527–39 (2010).

86. Zhang, S. C., Wernig, M., Duncan, I. D., Brustle, O. & Thomson, J. A. In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nat Biotechnol* **19,** 1129–33 (2001).

87. Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463,** 1035–41 (2010).

88. Zhang, Y. *et al.* Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* **78,** 785–98 (2013).

89. Marchetto, M. C. *et al.* Species-specific maturation profiles of human, chimpanzee and bonobo neural cells. *Elife* **8** (2019).

90. Lancaster, M. A. & Knoblich, J. A. Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* **345,** 1247125–1–1247125–9 (2014).

91. Spence, J. R. *et al.* Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* **470,** 105–9 (2011).

92. Lancaster, M. A. *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* **501,** 373–9 (2013).

93. Kadoshima, T. *et al.* Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex. *Proc Natl Acad Sci U S A* **110,** 20284–9 (2013).

94. Pasca, A. M. *et al.* Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nat Methods* **12,** 671–8 (2015).

95. Takebe, T. *et al.* Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* **499,** 481–4 (2013).

96. Takasato, M. *et al.* Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature* **526,** 564–8 (2015).

97. Dye, B. R. *et al.* In vitro generation of human pluripotent stem cell derived lung organoids. *Elife* **4** (2015).

98. Huang, L. *et al.* Ductal pancreatic cancer modeling and drug screening using human pluripotent stem cell- and patient-derived tumor organoids. *Nat Med* **21,** 1364–71 (2015).

99. Nakano, T. *et al.* Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell* **10,** 771–785 (2012).

100. Koehler, K. R. *et al.* Generation of inner ear organoids containing functional hair cells from human pluripotent stem cells. *Nat Biotechnol* **35,** 583–589 (2017).

101. Yumlu, S. *et al.* Gene editing and clonal isolation of human induced pluripotent stem cells using CRISPR/Cas9. *Methods* **121-122,** 29–44 (2017).

102. Li, Y. *et al.* Induction of expansion and folding in human cerebral organoids. *Cell Stem Cell* **20,** 385–396 (2017).

103. Eiraku, M. *et al.* Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell* **3,** 519–32 (2008).

104. Qian, X. *et al.* Brain-region-specific organoids using mini-bioreactors for modeling ZIKV exposure. *Cell* **165,** 1238–1254 (2016).

105. Jo, J. *et al.* Midbrain-like organoids from human pluripotent stem cells contain functional dopaminergic and neuromelanin-producing neurons. *Cell Stem Cell* **19,** 248–257 (2016).

106. Sakaguchi, H. *et al.* Generation of functional hippocampal neurons from self-organizing human embryonic stem cell-derived dorsomedial telencephalic tissue. *Nat Commun* **6,** 1–11 (2015).

107. Muguruma, K., Nishiyama, A., Kawakami, H., Hashimoto, K. & Sasai, Y. Self-organization of polarized cerebellar tissue in 3D culture of human pluripotent stem cells. *Cell Rep* **10,** 537–50 (2015).

108. Ozone, C. *et al.* Functional anterior pituitary generated in self-organizing culture of human embryonic stem cells. *Nat Commun* **7,** 10351–1–10351–10 (2016).

109. Birey, F. *et al.* Assembly of functionally integrated human forebrain spheroids. *Nature* **545,** 54–59 (2017).

110. Xiang, Y. *et al.* Fusion of regionally specified hPSC-derived organoids models human brain development and interneuron migration. *Cell Stem Cell* **21,** 383–398 (2017).

111. Bagley, J. A., Reumann, D., Bian, S., Levi-Strauss, J. & Knoblich, J. A. Fused cerebral organoids model interactions between brain regions. *Nat Methods* **14,** 743–751 (2017).

112. Xiang, Y. *et al.* hESC-derived thalamic organoids form reciprocal projections when fused with cortical organoids. *Cell Stem Cell* **24,** 487–497 (2019).

113. Renner, M. *et al.* Self-organized developmental patterning and differentiation in cerebral organoids. *EMBO J* **36,** 1316–1329 (2017).

114. Camp, J. G. *et al.* Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proc Natl Acad Sci U S A* **112,** 15672–7 (2015).

115. Luo, C. *et al.* Cerebral organoids recapitulate epigenomic signatures of the human fetal brain. *Cell Rep* **17,** 3369–3384 (2016).

116. Amiri, A. *et al.* Transcriptome and epigenome landscape of human cortical development modeled in organoids. *Science* **362,** 1–9 (2018).

117. Velasco, S. *et al.* Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature* **570,** 523–527 (2019).

118. Trujillo, C. A. *et al.* Complex oscillatory waves emerging from cortical organoids model early human brain network development. *Cell Stem Cell* **25,** 558–569 (2019).

119. Bershteyn, M. *et al.* Human iPSC-derived cerebral organoids model cellular features of lissencephaly and reveal prolonged mitosis of outer radial glia. *Cell Stem Cell* **20,** 435–449 (2017).

120. Klaus, J. *et al.* Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia. *Nat Med* **25,** 561–568 (2019).

121. Xu, R. *et al.* OLIG2 drives abnormal neurodevelopmental phenotypes in human iPSC-based organoid and chimeric mouse models of Down syndrome. *Cell Stem Cell* **24,** 908–926 (2019).

122. Mariani, J. *et al.* FOXG1-dependent dysregulation of GABA/glutamate neuron differentiation in autism spectrum disorders. *Cell* **162,** 375–390 (2015).

123. Garcez, P. P. *et al.* Zika virus impairs growth in human neurospheres and brain organoids. *Science* **352,** 816–8 (2016).

124. Cugola, F. R. *et al.* The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* **534,** 267–71 (2016).

125. Zhou, T. *et al.* High-content screening in hPSC-neural progenitors identifies drug candidates that inhibit Zika virus infection in fetal-like organoids and adult brain. *Cell Stem Cell* **21,** 274–283 (2017).

126. Raja, W. K. *et al.* Self-organizing 3D human beural tissue derived from induced pluripotent stem cells recapitulate Alzheimer's disease phenotypes. *PLoS One* **11,** e0161969 (2016).

127. Gonzalez, C. *et al.* Modeling amyloid beta and tau pathology in human cerebral organoids. *Mol Psychiatry* **23,** 2363–2374 (2018).

128. Quadrato, G. *et al.* Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* **545,** 48–53 (2017).

129. Lancaster, M. A. *et al.* Guided self-organization and cortical plate formation in human brain organoids. *Nat Biotechnol* **35,** 659–666 (2017).

130. Madhavan, M. *et al.* Induction of myelinating oligodendrocytes in human cortical spheroids. *Nat Methods* **15,** 700–706 (2018).

131. Marton, R. M. *et al.* Differentiation and maturation of oligodendrocytes in human three-dimensional neural cultures. *Nat Neurosci* **22,** 484–491 (2019).

132. Ormel, P. R. *et al.* Microglia innately develop within cerebral organoids. *Nat Commun* **9,** 4167 (2018).

133. Sloan, S. A. *et al.* Human astrocyte maturation captured in 3D cerebral cortical spheroids derived from pluripotent stem cells. *Neuron* **95,** 779–790 (2017).

134. Mansour, A. A. *et al.* An in vivo model of functional and vascularized human brain organoids. *Nat Biotechnol* **36,** 432–441 (2018).

135. Chimpanzee, S. & Analysis, C. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437,** 69–87 (2005).

136. King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188,** 107–16 (1975).

137. Haygood, R., Babbitt, C. C., Fedrigo, O. & Wray, G. A. Contrasts between adaptive coding and noncoding changes during human evolution. *Proc Natl Acad Sci U S A* **107,** 7853–7 (2010).

138. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478,** 483–9 (2011).

139. Somel, M. *et al.* MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. *PLoS Biol* **9,** e1001214 (2011).

140. Zhu, Y. *et al.* Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science* **362,** 1–15 (2018).

141. He, Z. *et al.* Comprehensive transcriptome analysis of neocortical layers in humans, chimpanzees and macaques. *Nat Neurosci* **20,** 886–895 (2017).

142. Oldham, M. C., Horvath, S. & Geschwind, D. H. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci U S A* **103,** 17973–8 (2006).

143. Nowick, K., Gernat, T., Almaas, E. & Stubbs, L. Differences in human and chimpanzee gene expression patterns define an evolving network of transcription factors in brain. *Proc Natl Acad Sci U S A* **106,** 22358–63 (2009).

144. Konopka, G. *et al.* Human-specific transcriptional networks in the brain. *Neuron* **75,** 601–17 (2012).

145. Berto, S. & Nowick, K. Species-specific changes in a primate transcription factor network provide insights into the molecular evolution of the primate prefrontal cortex. *Genome Biol Evol* **10,** 2023–2036 (2018).

146. Caceres, M. *et al.* Elevated gene expression levels distinguish human from non-human primate brains. *Proc Natl Acad Sci U S A* **100,** 13030–5 (2003).

147. Khaitovich, P. *et al.* A neutral model of transcriptome evolution. *PLoS Biol* **2,** e132 (2004).

148. Khaitovich, P. *et al.* Regional patterns of gene expression in human and chimpanzee brains. *Genome Res* **14,** 1462–73 (2004).

149. Khaitovich, P. *et al.* Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309,** 1850–4 (2005).

150. Khaitovich, P. *et al.* Positive selection on gene expression in the human brain. *Curr Biol* **16,** 356–8 (2006).

151. Reilly, S. K. *et al.* Evolutionary changes in promoter and enhancer activity during human corticogenesis. *Science* **347,** 1155–9 (2015).

152. Vermunt, M. W. *et al.* Epigenomic annotation of gene regulatory alterations during evolution of the primate brain. *Nat Neurosci* **19,** 494–503 (2016).

153. Shulha, H. P. *et al.* Human-specific histone methylation signatures at transcription start sites in prefrontal neurons. *PLoS Biol* **10,** e1001427 (2012).

154. Prescott, S. L. *et al.* Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell* **163,** 68–83 (2015).

155. Hu, H. Y. *et al.* MicroRNA expression and regulation in human, chimpanzee, and macaque brains. *PLoS Genet* **7,** e1002327 (2011).

156. He, Z., Bammann, H., Han, D., Xie, G. & Khaitovich, P. Conserved expression of lincRNA during human and macaque prefrontal cortex development and maturation. *RNA* **20,** 1103–11 (2014).

157. Zeng, J. *et al.* Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution. *Am J Hum Genet* **91,** 455–65 (2012).

158. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6,** 377–82 (2009).

159. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573,** 61–68 (2019).

160. Wu, A. R. *et al.* Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods* **11,** 41–6 (2014).

161. Treutlein, B. *et al.* Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509,** 371–5 (2014).

162. Pollen, A. A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat Biotechnol* **32,** 1053–8 (2014).

163. Jaitin, D. A. *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343,** 776–9 (2014).

164. Gierahn, T. M. *et al.* Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods* **14,** 395–398 (2017).

165. Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161,** 1187–1201 (2015).

166. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161,** 1202–1214 (2015).

167. Rosenberg, A. B. *et al.* Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360,** 176–182 (2018).

168. Gerber, T. *et al.* Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. *Science* **362** (2018).

169. Shalek, A. K. *et al.* Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* **498,** 236–40 (2013).

170. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348,** aaa6090–1–aaa6090–14 (2015).

171. Stahl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353,** 78–82 (2016).

172. Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361,** 1–9 (2018).

173. Eng, C. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568,** 235–239 (2019).

174. Biddy, B. A. *et al.* Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564,** 219–224 (2018).

175. Spanjaard, B. *et al.* Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat Biotechnol* **36,** 469–473 (2018).

176. Dixit, A. *et al.* Perturb-seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167,** 1853–1866 (2016).

177. Datlinger, P. *et al.* Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods* **14,** 297–301 (2017).

178. Tabula Muris, C. *et al.* Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562,** 367–372 (2018).

179. Regev, A. *et al.* The human cell atlas. *Elife* **6,** 1–30 (2017).

180. Muraro, M. J. *et al.* A single-cell transcriptome atlas of the human pancreas. *Cell Syst* **3,** 385–394 e3 (2016).

181. Villani, A. C. *et al.* Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356** (2017).

182. Aizarani, N. *et al.* A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572,** 199–204 (2019).

183. Lukowski, S. W. *et al.* A single-cell transcriptome atlas of the adult human retina. *EMBO J* **38,** e100811–1–e100811–15 (2019).

184. Vieira Braga, F. A. *et al.* A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med* **25,** 1153–1163 (2019).

185. Venteicher, A. S. *et al.* Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science* **355,** 1–11 (2017).

186. Chung, W. *et al.* Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun* **8,** 1–12 (2017).

187. Velmeshev, D. *et al.* Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364,** 685–689 (2019).

188. Schirmer, L. *et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573,** 75–82 (2019).

189. Farrell, J. A. *et al.* Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360,** 1–7 (2018).

190. Karaiskos, N. *et al.* The Drosophila embryo at single-cell transcriptome resolution. *Science* **358,** 194–199 (2017).

191. Packer, J. S. *et al.* A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. *Science* **365,** 1–8 (2019).

192. Aztekin, C. *et al.* Identification of a regeneration-organizing cell in the Xenopus tail. *Science* **364,** 653–658 (2019).

193. Plass, M. *et al.* Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science* **360,** 1–10 (2018).

194. Ramskold, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* **30,** 777–82 (2012).

195. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* **10,** 1096–8 (2013).

196. Navin, N. *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **472,** 90–4 (2011).

197. Buenrostro, J. D. *et al.* Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523,** 486–90 (2015).

198. Cusanovich, D. A. *et al.* Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348,** 910–4 (2015).

199. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol* **19,** 161 (2018).

200. Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502,** 59–64 (2013).

201. Ramani, V. *et al.* Massively multiplex single-cell Hi-C. *Nat Methods* **14,** 263–266 (2017).

202. Macaulay, I. C. *et al.* G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods* **12,** 519–22 (2015).

203. Dey, S. S., Kester, L., Spanjaard, B., Bienko, M. & van Oudenaarden, A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol* **33,** 285–289 (2015).

204. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* **14,** 865–868 (2017).

205. Cao, J. *et al.* Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361,** 1380–1385 (2018).

206. Clark, S. J. *et al.* scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun* **9,** 1–9 (2018).

207. Li, G. *et al.* Joint profiling of DNA methylation and chromatin architecture in single cells. *Nat Methods* **16,** 991–993 (2019).

208. Zeisel, A. *et al.* Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347,** 1138–42 (2015).

209. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci* **19,** 335–46 (2016).

210. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563,** 72–78 (2018).

211. Gouwens, N. W. *et al.* Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat Neurosci* **22,** 1182–1195 (2019).

212. Cadwell, C. R. *et al.* Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nat Biotechnol* **34,** 199–203 (2016).

213. Mayer, S. *et al.* Multimodal single-cell analysis reveals physiological maturation in the developing human neocortex. *Neuron* **102,** 143–158 (2019).

214. Grindberg, R. V. *et al.* RNA-sequencing from single nuclei. *Proc Natl Acad Sci U S A* **110,** 19802–7 (2013).

215. Lake, B. B. *et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352,** 1586–90 (2016).

216. Bakken, T. E. *et al.* Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS One* **13,** e0209648 (2018).

217. Johnson, M. B. *et al.* Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat Neurosci* **18,** 637–46 (2015).

218. Darmanis, S. *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci U S A* **112,** 7285–90 (2015).

219. Nowakowski, T. J. *et al.* Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358,** 1318–1323 (2017).

220. Zhong, S. *et al.* A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555,** 524–528 (2018).

221. Lake, B. B. *et al.* Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat Biotechnol* **36,** 70–80 (2018).

222. Mora-Bermudez, F. *et al.* Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development. *Elife* **5** (2016).

223. Sousa, A. M. M., Meyer, K. A., Santpere, G., Gulden, F. O. & Sestan, N. Evolution of the human nervous system function, structure, and development. *Cell* **170,** 226–247 (2017).

224. Zheng, G. X. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat Commun* **8,** 14049 (2017).

225. Angerer, P. *et al.* destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32,** 1241–3 (2016).

226. Aghakhani, Y. *et al.* The role of periventricular nodular heterotopia in epileptogenesis. *Brain* **128,** 641–51 (2005).

227. Mansour, S. *et al.* Van Maldergem syndrome: further characterisation and evidence for neuronal migration abnormalities and autosomal recessive inheritance. *Eur J Hum Genet* **20,** 1024–31 (2012).

228. Pacary, E. *et al.* Proneural transcription factors regulate different steps of cortical neuron migration through Rnd-mediated inhibition of RhoA signaling. *Neuron* **69,** 1069–84 (2011).

229. Xie, S., Duan, J., Li, B., Zhou, P. & Hon, G. C. Multiplexed engineering and analysis of combinatorial enhancer activity in single cells. *Mol Cell* **66,** 285–299 (2017).

230. Raj, B. *et al.* Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat Biotechnol* **36,** 442–450 (2018).

231. Ludwig, L. S. *et al.* Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* **176,** 1325–1339 (2019).

232. Faridani, O. R. *et al.* Single-cell sequencing of the small-RNA transcriptome. *Nat Biotechnol* **34,** 1264–1266 (2016).

233. Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5,** 877–9 (2008).

234. Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat Methods* **15,** 932–935 (2018).

235. Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci* **10,** 724–35 (2009).

236. Striedter, G. Principles of Brain Evolution (2005).

237. Borrell, V. & Reillo, I. Emerging roles of neural stem cells in cerebral cortex development and evolution. *Dev Neurobiol* **72,** 955–71 (2012).

238. Florio, M. & Huttner, W. B. Neural progenitors, neurogenesis and the evolution of the neocortex. *Development* **141,** 2182–94 (2014).

239. Herculano-Houzel, S. The human brain in numbers: a linearly scaled-up primate brain. *Front Hum Neurosci* **3,** 1–11 (2009).

240. Lui, J. H., Hansen, D. V. & Kriegstein, A. R. Development and evolution of the human neocortex. *Cell* **146,** 18–36 (2011).

241. Lewitus, E., Kelava, I. & Huttner, W. B. Conical expansion of the outer subventricular zone and the role of neocortical folding in evolution and development. *Front Hum Neurosci* **7,** 424 (2013).

242. Angevine, J. Embryonic vertebrate central nervous system: revised terminology. The Boulder Committee. *Anat Rec* **166,** 257–61 (1970).

243. Dehay, C., Kennedy, H. & Kosik, K. S. The outer subventricular zone and primate-specific cortical complexification. *Neuron* **85,** 683–94 (2015).

244. Götz, M. & Huttner, W. B. The cell biology of neurogenesis. *Nat Rev Mol Cell Biol* **6,** 777–88 (2005).

245. Lancaster, M. A. & Knoblich, J. A. Spindle orientation in mammalian cerebral cortical development. *Curr Opin Neurobiol* **22,** 737–46 (2012).

246. Mora-Bermudez, F. & Huttner, W. B. Novel insights into mammalian embryonic neural stem cell division: focus on microtubules. *Mol Biol Cell* **26,** 4302–6 (2015).

247. Mora-Bermudez, F., Matsuzaki, F. & Huttner, W. B. Specific polar subpopulations of astral microtubules control spindle orientation and symmetric neural stem cell division. *Elife* **3** (2014).

248. Shitamukai, A. & Matsuzaki, F. Control of asymmetric cell division of mammalian neural progenitors. *Dev Growth Differ* **54,** 277–86 (2012).

249. Kriegstein, A. & Alvarez-Buylla, A. The glial nature of embryonic and adult neural stem cells. *Annu Rev Neurosci* **32,** 149–84 (2009).

250. Lancaster, M. A. & Knoblich, J. A. Generation of cerebral organoids from human pluripotent stem cells. *Nat Protoc* **9,** 2329–40 (2014).

251. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32,** 381–386 (2014).

252. Iacopetti, P. *et al.* Expression of the antiproliferative gene TIS21 at the onset of neurogenesis identifies single neuroepithelial cells that switch from proliferative to neuron-generating division. *Proc Natl Acad Sci U S A* **96,** 4639–44 (1999).

253. Pilaz, L. J. *et al.* Prolonged mitosis of neural progenitors alters cell fate in the developing brain. *Neuron* **89,** 83–99 (2016).

254. Arai, Y. *et al.* Neural stem and progenitor cells shorten S-phase on commitment to neuron production. *Nat Commun* **2,** 154 (2011).

255. Lewitus, E., Kelava, I., Kalinka, A. T., Tomancak, P. & Huttner, W. B. An adaptive threshold in mammalian neocortical evolution. *PLoS Biol* **12,** e1002000 (2014).

256. Musacchio, A. & Salmon, E. D. The spindle-assembly checkpoint in space and time. *Nat Rev Mol Cell Biol* **8,** 379–93 (2007).

257. Konno, D. *et al.* Neuroepithelial progenitors undergo LGN-dependent planar divisions to maintain self-renewability during mammalian neurogenesis. *Nat Cell Biol* **10,** 93–101 (2008).

258. Kosodo, Y. *et al.* Asymmetric distribution of the apical plasma membrane during neurogenic divisions of mammalian neuroepithelial cells. *EMBO J* **23,** 2314–24 (2004).

259. Tosato, G. & Cohen, J. I. Generation of Epstein-Barr Virus (EBV)-immortalized B cell lines. *Curr Protoc Immunol* **Chapter 7,** Unit 7 22 (2007).

260. Renaud, G., Kircher, M., Stenzel, U. & Kelso, J. freeIbis: an efficient basecaller with calibrated quality scores for Illumina sequencers. *Bioinformatics* **29,** 1208–9 (2013).

261. Renaud, G., Stenzel, U. & Kelso, J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res* **42,** e141 (2014).

262. Renaud, G., Stenzel, U., Maricic, T., Wiebe, V. & Kelso, J. deML: robust demultiplexing of Illumina sequences using a likelihood-based approach. *Bioinformatics* **31,** 770–2 (2015).

263. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7,** 562–78 (2012).

264. Team, R. D. C. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria (2010).

265. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4,** 44–57 (2009).

266. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).

267. Anders, S., Pyl, P. T. & Huber, W. HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–9 (2015).

268. Kharchenko, P. V., Silberstein, L. & Scadden, D. T. Bayesian approach to single-cell differential expression analysis. *Nat Methods* **11,** 740–2 (2014).

269. Scialdone, A. *et al.* Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods* **85,** 54–61 (2015).

270. Nowakowski, R. S., Lewin, S. B. & Miller, M. W. Bromodeoxyuridine immunohistochemical determination of the lengths of the cell cycle and the DNA-synthetic phase for an anatomically defined population. *J Neurocytol* **18,** 311–8 (1989).

271. Mora-Bermudez, F. & Ellenberg, J. Measuring structural dynamics of chromosomes in living cells by fluorescence microscopy. *Methods* **41,** 158–67 (2007).

272. Ma, H. *et al.* Abnormalities in human pluripotent cells due to reprogramming mechanisms. *Nature* **511,** 177–83 (2014).

273. Martin, R. D. Relative brain size and basal metabolic rate in terrestrial vertebrates. *Nature* **293,** 57–60 (1981).

274. Weinreb, C., Wolock, S. & Klein, A. M. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *Bioinformatics* **34,** 1246–1248 (2018).

275. La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560,** 494–498 (2018).

276. Leigh, S. R. Brain growth, life history, and cognition in primate and human evolution. *Am J Primatol* **62,** 139–64 (2004).

277. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser–a database of tissue-specific human enhancers. *Nucleic Acids Res* **35,** 88–92 (2007).

278. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* **14,** 975–978 (2017).

279. Gregorio, I., Braghetta, P., Bonaldo, P. & Cescon, M. Collagen VI in healthy and diseased nervous system. *Dis Model Mech* **11** (2018).

280. Kilpinen, H. *et al.* Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* **546,** 370–375 (2017).

281. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282,** 1145–7 (1998).

282. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol* **36,** 89–94 (2018).

283. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36,** 411–420 (2018).

284. Dougherty, M. L. *et al.* Transcriptional fates of human-specific segmental duplications in brain. *Genome Res* **28,** 1566–1576 (2018).

285. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10,** 1213–8 (2013).

286. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–9 (2009).

287. Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).

288. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9,** R137 (2008).

289. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29,** 24–6 (2011).

290. Pliner, H. A. *et al.* Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. *Mol Cell* **71,** 858–871 (2018).

291. Preissl, S. *et al.* Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat Neurosci* **21,** 432–439 (2018).

292. Suzuki, K. *et al.* Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell Stem Cell* **15,** 31–6 (2014).

293. Mayer, C. *et al.* Developmental diversification of cortical inhibitory interneurons. *Nature* **555,** 457–462 (2018).

294. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28,** 495–501 (2010).

295. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–2 (2010).

296. Lindblad-Toh, K. *et al.* A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478,** 476–82 (2011).

297. Prabhakar, S., Noonan, J. P., Paabo, S. & Rubin, E. M. Accelerated evolution of conserved noncoding sequences in humans. *Science* **314,** 786 (2006).

298. Gittelman, R. M. *et al.* Comprehensive identification and analysis of human accelerated regulatory DNA. *Genome Res* **25,** 1245–55 (2015).

299. Cagan, A. *et al.* Natural selection in the great apes. *Mol Biol Evol* **33,** 3268–3283 (2016).

300. Peyregne, S., Boyle, M. J., Dannemann, M. & Prüfer, K. Detecting ancient positive selection in humans using extended lineage sorting. *Genome Res* **27,** 1563–1572 (2017).

301. Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505,** 43–9 (2014).

302. Chintalapati, M., Dannemann, M. & Prüfer, K. Using the Neandertal genome to study the evolution of small insertions and deletions in modern humans. *BMC Evol Biol* **17,** 179 (2017).

303. Fu, Y. *et al.* FunSeq2: a framework for prioritizing noncoding regulatory variants in cancer. *Genome Biol* **15,** 480 (2014).

304. Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* **172,** 650–665 (2018).

305. Yu, Q. & He, Z. Comprehensive investigation of temporal and autism-associated cell type composition-dependent and independent gene expression changes in human brains. *Sci Rep* **7,** 4121 (2017).

306. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20,** 296 (2019).

307. Tekinay, A. B. *et al.* A role for LYNX2 in anxiety-related behavior. *Proc Natl Acad Sci U S A* **106,** 4477–82 (2009).

308. Reijnders, M. R. F. *et al.* RAC1 missense mutations in developmental disorders with diverse phenotypes. *Am J Hum Genet* **101,** 466–477 (2017).

309. Nakamura, M. T. & Nara, T. Y. Structure, function, and dietary regulation of delta6, delta5, and delta9 desaturases. *Annu Rev Nutr* **24,** 345–76 (2004).

310. Cho, H. P., Nakamura, M. & Clarke, S. D. Cloning, expression, and fatty acid regulation of the human delta-5 desaturase. *J Biol Chem* **274,** 37335–9 (1999).

311. Di Gregorio, M. C. *et al.* pH sensitive tubules of a bile acid derivative: a tubule opening by release of wall leaves. *Phys Chem Chem Phys* **15,** 7560–6 (2013).

312. Halevi, S. *et al.* Conservation within the RIC-3 gene family. Effectors of mammalian nicotinic acetylcholine receptor expression. *J Biol Chem* **278,** 34411–7 (2003).

313. Romero, D. M., Bahi-Buisson, N. & Francis, F. Genetics and mechanisms leading to human cortical malformations. *Semin Cell Dev Biol* **76,** 33–75 (2018).

314. Liu, J. S. Molecular genetics of neuronal migration disorders. *Curr Neurol Neurosci Rep* **11,** 171–8 (2011).

315. Cardoso, C. *et al.* Periventricular heterotopia, mental retardation, and epilepsy associated with 5q14.3-q15 deletion. *Neurology* **72,** 784–92 (2009).

316. Dubeau, F. *et al.* Periventricular and subcortical nodular heterotopia. A study of 33 patients. *Brain* **118,** 1273–87 (1995).

317. Tassi, L. *et al.* Electroclinical, MRI and neuropathological study of 10 patients with nodular heterotopia, with surgical outcomes. *Brain* **128,** 321–37 (2005).

318. Heinzen, E. L. *et al.* De novo and inherited private variants in MAP1B in periventricular nodular heterotopia. *PLoS Genet* **14,** e1007281 (2018).

319. O'Neill, A. C. *et al.* A primate-specific isoform of PLEKHG6 regulates neurogenesis and neuronal migration. *Cell Rep* **25,** 2729–2741 (2018).

320. Ishiuchi, T., Misaki, K., Yonemura, S., Takeichi, M. & Tanoue, T. Mammalian Fat and Dachsous cadherins regulate apical membrane organization in the embryonic cerebral cortex. *J Cell Biol* **185,** 959–67 (2009).

321. Cappello, S. *et al.* The Rho-GTPase cdc42 regulates neural progenitor fate at the apical surface. *Nat Neurosci* **9,** 1099–107 (2006).

322. Topol, A., Tran, N. N. & Brennand, K. J. A guide to generating and using hiPSC derived NPCs for the study of neurological diseases. *J Vis Exp,* e52495 (2015).

323. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14,** 979–982 (2017).

324. Wang, J. *et al.* Epilepsy-associated genes. *Seizure* **44,** 11–20 (2017).

325. Riesenberg, S. & Maricic, T. Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells. *Nat Commun* **9,** 2164 (2018).

326. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* (2010).

327. Boyer, L. F., Campbell, B., Larkin, S., Mu, Y. & Gage, F. H. Dopaminergic differentiation of human pluripotent cells. *Curr Protoc Stem Cell Biol* (2012).

328. Pilz, G. A. *et al.* Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type. *Nat Commun* **4,** 2125 (2013).

329. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–11 (2009).

330. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9,** 357–9 (2012).

331. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28,** 511–5 (2010).

Acknowledgments

When I started my PhD almost five years ago, I could only guess what kind of journey was ahead of me, both on a personal and professional level. Many people were directly or indirectly part of this undertaking as supervisors, colleagues, cooperation partners, friends and family, and it is almost impossible to do justice to all of them here.

First and foremost, I would like to thank my PhD supervisor, Barbara Treutlein, for giving me the opportunity and the trust to join her new group in the still young field of single-cell genomics, and to start this exciting scientific journey together. With her positive and energetic attitude, Barbara has created a special atmosphere in the Single Cell Genomics group which has been extremely motivating, productive and supportive. I am grateful that I got the opportunity to work on interdisciplinary and cooperative projects in the course of my PhD from which I have learned so much.

Furthermore, I would like to thank all the current and former members of the Single Cell Genomics group for being part of this inspiring environment. Above all, I would like to thank Gray Camp, without whom the Single Cell Genomics group would certainly not be as it is.

I would also like to thank Svante Pääbo who creates a unique working atmosphere in the Department of Evolutionary Genetics, where I felt very welcome and recognized from the first day of my PhD time and I am grateful that I could be part of this special group of people. I would also like to thank him and all members of the Thursday group for insightful discussions.

I am in particular grateful to a number of people, both from the Single Cell Genomics Group and from the Department of Evolutionary Genetics, who made this time the enriching time, it has become for me: Tobias, my fellow PhD student from day one for scientific and non-scientific discussions; Agnieska for sharing the office and the ups and downs of the scientific work endeavour; Damian for great discussions during lunch, coffee breaks and other opportunities; Jorge for lively discussions about science and spontaneous visits in Berlin; Zhisong for sharing the office with me and for scientific discussions; Mike for sharing the organoid experience; Marie for rememberable conferences; Fatima for sharing the passion for science and music; Gosia for support in and outside the lab; Anne, Theresa and Kathrin for making life in the cell culture going smoothly; Birgit for making the lab a well-organized place; Dominik for being my concert companion; Antje and Birgit for the great sequencing support; Ben Vernot for insights on analyses; and Rebecca Snabel for her work as a Master student. Also many thanks to Viola for making bureaucratic life so much easier and Ines and Rigo for the smooth IT support.

Moreover, many thanks to the Boehringer Ingelheim Fonds for supporting me for two years and to the inspiring people that I met at the seminars and the annual Christmas Party in Mainz.

Furthermore, I would like to thank our collaborators with whom we have achieved new interesting scientific insights and had fruitful collaborations during the last years: Wieland Huttner and members from his group at the MPI CBG in Dresden, specifically Farhath Badsha, Michael Heide, Felipe Mora-Bermúdez, Lea Berninger and Jula Peters; Silvia Cappello and members from her group at the MPI of Psychiatry in Munich, especially Johannes Klaus and Christina Kyrousi; and Philipp Khaitovich and members from his group at the MPI CAS in Shanghai, specifically Patricia Guijarro, Dingding Han and Zhengzong Qian.

Außerdem möchte ich meinen Eltern und meinen Großeltern danken, die mich in meinen Entscheidungen immer unterstützt haben und ohne die ich sicher nicht dort wäre, wo ich jetzt bin, sowohl vor als auch während der Doktorarbeit.

Vielen Dank an Franziska - auch aus eigener Erfahrung weißt du wahrscheinlich am besten, was die PhD-Zeit für mich bedeutet und ich bin dankbar dafür, dass wir einen Großteil dieser Zeit gemeinsam gemeistert haben. Außerdem möchte ich meinen Freunden aus Schul-, Studien- und Doktorandenzeit, vor allem Michel, Carolin, Jana, Madlen, Heni, Tina, Carolann, Agneta, Anni und Renata, die mich während der letzten Jahre begleitet haben, danken.

Abschließend möchte ich mich bei Prof. Wolfgang Enard für die Übernahme des Zweitgutachtens bedanken.

Declaration of Independence

I hereby declare that I have written and conceived this thesis without any inadmissible help and without using materials other than those explicitly stated. I have not previously attempted to complete this or any other doctorate degree.

Leipzig, November 19th, 2019 Sabina Kanton

Author contribution statement, Sabina Kanton
Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development

---

Author contribution statement

Title: **Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development**

Journal: Elife

Authors: Felipe Mora-Bermúdez*, Farhath Badsha*, Sabina Kanton*, J. Gray Camp*, Benjamin Vernot, Kathrin Köhler, Birger Voigt, Keisuke Okita, Tomislav Maricic, Zhisong He, Robert Lachmann, Svante Pääbo, Barbara Treutlein, Wieland B. Huttner

---

Part Felipe Mora-Bermúdez

- Conceived the study
- Designed experiments
- Performed and analyzed time lapse experiments
- Wrote the manuscript

Part Farhath Badsha

- Conceived the study
- Designed experiments
- Generated cerebral organoids
- Performed immunohistochemical stainings and EdU labeling on organoids
- Analyzed immunohistochemical stainings and EdU labeling on organoids
- Wrote the manuscript

Part Sabina Kanton

- Conceived the study
- Designed experiments
- Generated cerebral organoids
- Performed single cell RNA-Seq experiments and generated single cell sequencing data
- Analyzed single cell RNA-Seq data
- Wrote the manuscript

Part J. Gray Camp

- Conceived the study
- Designed experiments
- Performed single cell RNA-Seq experiments and generated single cell sequencing data
- Analyzed single cell RNA-Seq data
- Wrote the manuscript

Part Benjamin Vernot

- Analyzed single cell RNA-Seq data
- Assisted in interpreting single cell RNA-Seq data

Part Kathrin Köhler

- Prepared induced pluripotent stem cell lines SandraA and Toba
- Assisted in iPS cell culture
- Provided information relevant for interpretation of data

Part Birger Voigt

- Prepared induced pluripotent stem cell lines SandraA and Toba
- Provided information relevant for interpretation of data

Part Keisuke Okita

- Prepared induced pluripotent stem cell lines SandraA and Toba
- Provided information relevant for interpretation of data

Part Tomislav Maricic

- Prepared induced pluripotent stem cell lines SandraA and Toba
- Provided information relevant for interpretation of data

Part Zhisong He

- Provided consensus genome

Part Robert Lachmann

- Provided human fetal tissue
- Provided information relevant for interpretation of data

Part Svante Päabo

- Conceived the study
- Designed experiments
- Provided intellectual guidance in the interpretation of the data
- Wrote the manuscript

Part Barbara Treutlein

- Conceived the study
- Designed experiments
- Analyzed single cell RNA-Seq data
- Provided intellectual guidance in the interpretation of the data
- Wrote the manuscript

Part Wieland B. Huttner

- Conceived the study
- Designed experiments
- Provided intellectual guidance in the interpretation of the data
- Wrote the manuscript

S. Kanton

Sabina Kanton

Wieland Huttner

Wieland B. Huttner

S. Kanton

Sabina Kanton

Svante Pääbo

B. Treutlein

Barbara Treutlein

Author contribution statement, Sabina Kanton
Organoid single-cell genomic atlas uncovers human-specific features of brain development

---

Author contribution statement

Title: **Organoid single-cell genomic atlas uncovers human-specific features of brain development**

Journal: Nature

Authors: Sabina Kanton*, Michael J. Boyle*, Zhisong He*, Malgorzata Santel, Anne Weigert, Fatima Sanchis Calleja, Patricia Guijarro, Leila Sidow, Jonas Fleck, Dingding Han, Zhengzong Qian, Michael Heide, Wieland Huttner, Philipp Khaitovich, Svante Pääbo, Barbara Treutlein, J. Gray Camp

---

Part Sabina Kanton

- Designed the study
- Generated cerebral organoids
- Performed single cell and single nucleus RNA-Seq experiments and generated sequencing data
- Analyzed single cell and single nucleus RNA-Seq data
- Wrote the manuscript

Part Michael J Boyle

- Designed the study
- Generated cerebral organoids
- Performed single cell ATAC-Seq experiments and generated sequencing data
- Analyzed single cell ATAC-Seq data
- Wrote the manuscript

Part Zhisong He

- Designed the study
- Analyzed single cell RNA-Seq, single nucleus RNA-Seq and ATAC-Seq data
- Wrote the manuscript

Part Malgorzata Santel

- Performed sorting experiments
- Assisted in generating single cell and single nucleus RNA-Seq data

Part Anne Weigert

- Assisted in generating cerebral organoids
- Assisted in induced pluripotent stem cell culture

Part Fatima Sanchis Calleja

- Performed immunohistochemical stainings

Part Patricia Guijarro

- Dissected and sliced cubes from frozen brain tissue

Part Leila Sidow

- Assisted in generating cerebral organoids

Part Jonas Fleck

- Compared organoid single cell RNA-Seq data to mouse voxel maps

Part Dingding Han

- Generated bulk RNA-Seq libraries from frozen brain tissue

Part Zhengzong Qian

- Generated bulk RNA-Seq libraries from frozen brain tissue

Part Michael Heide

- Assisted in generating cerebral organoids
- Performed immunohistochemical stainings

Part Wieland B. Huttner

- Provided intellectual guidance

Part Philipp Khaitovich

- Provided intellectual guidance

Part Svante Pääbo

- Provided intellectual guidance

Part Barbara Treutlein

- Supervised the study
- Designed the study
- Wrote the manuscript

Part J. Gray Camp

- Supervised the study
- Designed the study
- Wrote the manuscript

S. Kanton

Sabina Kanton

16/9/2019

J. Gray Camp

Part Patricia Guijarro

- Dissected and sliced cubes from frozen brain tissue

Part Leila Sidow

- Assisted in generating cerebral organoids

Part Jonas Fleck

- Compared organoid single cell RNA-Seq data to mouse voxel maps

Part Dingding Han

- Generated bulk RNA-Seq libraries from frozen brain tissue

Part Zhengzong Qian

- Generated bulk RNA-Seq libraries from frozen brain tissue

Part Michael Heide

- Assisted in generating cerebral organoids
- Performed immunohistochemical stainings

Part Wieland B. Huttner

- Provided intellectual guidance

Part Philipp Khaitovich

- Provided intellectual guidance

Part Svante Pääbo

- Provided intellectual guidance

Part Barbara Treutlein

- Supervised the study
- Designed the study
- Wrote the manuscript

Part J. Gray Camp

- Supervised the study
- Designed the study
- Wrote the manuscript

S. Kanton
Sabina Kanton

B. Treutlein
Barbara Treutlein

Author contribution statement, Sabina Kanton
Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia

Author contribution statement

Title: **Altered neuronal migratory trajectories in human cerebral organoids derived from individuals with neuronal heterotopia**

Journal: Nature Medicine

Authors: Johannes Klaus*, Sabina Kanton*, Christina Kyrousi*, Ane Cristina Ayo-Martin, Rossella Di Giaimo, Stephan Riesenberg, Adam C. O'Neill, J. Gray Camp, Chiara Tocco, Malgorzata Santel, Ejona Rusha, Micha Drukker, Mariana Schroeder, Magdalena Götz, Stephen P. Robertson, Barbara Treutlein, Silvia Cappello

Part Johannes Klaus

- Generated cerebral organoids
- Performed immunohistochemical analysis of organoids
- Performed time lapse experiments
- Analyzed immunohistochemical and time lapse data
- Wrote the manuscript

Part Sabina Kanton

- Performed single cell RNA-Seq experiments and generated single cell sequencing data
- Analyzed single cell RNA-Seq data
- Assisted in generating cerebral organoids
- Wrote the manuscript

Part Christina Kyrousi

- Generated cerebral organoids
- Performed immunohistochemical analysis of organoids
- Analyzed immunohistochemical data
- Wrote the manuscript

Part Ane Cristina Ayo-Martin

- Generated cerebral organoids
- Performed immunohistochemical analysis of organoids

Part Rossella Di Giaimo

- Performed FACS analysis

Part Stephan Riesenberg

- Generated knock out induced pluripotent stem cell lines

Part Adam C. O'Neill

- Performed bioinformatic analysis of patients with periventricular heterotopia

Part J. Gray Camp

- Analyzed single cell RNA-Seq data

Part Chiara Tocco

- Performed immunohistochemical analysis of organoids

Part Malgorzata Santel

- Performed sorting experiments
- Assisted in generating single cell RNA-Seq data

Part Ejona Rusha

- Generated induced pluripotent stem cell lines

Part Micha Drukker

- Generated induced pluripotent stem cell lines

Part Mariana Schroeder

- Performed cluster analysis on time lapse imaging data

Part Magdalena Götz

- Assisted in supervising the project
- Assisted in data interpretation

Part Stephen P. Robertson

- Collected patient samples

Part Barbara Treutlein

- Supervised the project
- Analyzed single cell RNA-Seq data
- Wrote the manuscript

Part Silvia Cappello

- Conceived the study
- Designed the study
- Supervised the project
- Wrote the manuscript

Sabina Kanton

Silvia Cappello

Part Adam C. O'Neill

- Performed bioinformatic analysis of patients with periventricular heterotopia

Part J. Gray Camp

- Analyzed single cell RNA-Seq data

Part Chiara Tocco

- Performed immunohistochemical analysis of organoids

Part Malgorzata Santel

- Performed sorting experiments
- Assisted in generating single cell RNA-Seq data

Part Ejona Rusha

- Generated induced pluripotent stem cell lines

Part Micha Drukker

- Generated induced pluripotent stem cell lines

Part Mariana Schroeder

- Performed cluster analysis on time lapse imaging data

Part Magdalena Götz

- Assisted in supervising the project
- Assisted in data interpretation

Part Stephen P. Robertson

- Collected patient samples

Part Barbara Treutlein

- Supervised the project
- Analyzed single cell RNA-Seq data
- Wrote the manuscript

Part Silvia Cappello

- Conceived the study
- Designed the study
- Supervised the project
- Wrote the manuscript

Sabina Kanton

Barbara Treutlein